

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C07K 14/00</b>		<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 99/47552</b>
			<b>(43) International Publication Date:</b> 23 September 1999 (23.09.99)
<b>(21) International Application Number:</b> PCT/EP99/01779 <b>(22) International Filing Date:</b> 17 March 1999 (17.03.99)  <b>(30) Priority Data:</b> 09/042,763 17 March 1998 (17.03.98) US  <b>(71) Applicant (for all designated States except AT US):</b> NOVARTIS AG [CH/CH]; Schwarzwaldallee 215, CH-4058 Basel (CH).  <b>(71) Applicant (for AT only):</b> NOVARTIS-ERFINDUNGEN VERWALTUNGSGESELLSCHAFT M.B.H. [AT/AT]; Brunner Strasse 59, A-1235 Vienna (AT).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> SALMERON, John, Manuel [US/US]; 1308 Blackberry Lane, Hillsborough, NC 27278 (US). WEISLO, Laura, Jean [US/US]; 914 West South Street, Raleigh, NC 27603 (US). STRAWN, Leslie, Jean [US/US]; 32 Damsey Circle, Durham, NC 27713 (US). KRAMER, Catherine, Mae [US/US]; 608 Dana Court, Hillsborough, NC 27278 (US).  <b>(74) Agent:</b> BECKER, Konrad; Novartis AG, Corporate Intellectual Property, Patent & Trademark Dept., CH-4002 Basel (CH).			<b>(81) Designated States:</b> AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
<b>(54) Title:</b> GENES CONTROLLING DISEASES			
<b>(57) Abstract</b>  This invention describes genes encoding proteins which control resistance of plants to fungal pathogens. The invention also describes transgenic plants resistant to fungal pathogens and methods for making plants resistant to fungal pathogens. The invention further discloses a method to isolate additional genes coding for additional proteins controlling the resistance of plants to fungal pathogens.			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## GENES CONTROLLING DISEASES

The invention describes nucleotide sequences encoding proteins controlling resistance of plants to fungal diseases. The invention also relates to plants resistant to fungal diseases, and to methods of making plants resistant to fungal diseases.

Fungal diseases are responsible for yearly losses of about \$ 9.1 bn on agricultural crops in the US, and are caused by a large variety of biologically diverse pathogens. Different strategies have traditionally been used to control them. Resistance traits have been bred into agriculturally important varieties, thus providing various levels of resistance against either a narrow range of pathogen isolates or races, or against a broader range. However, this involves the long and labor intensive process of introducing desirable traits into commercial lines by genetic crosses and, due to the risk of pests evolving to overcome natural plant resistance, a constant effort to breed new resistance traits into commercial lines is required. Alternatively, fungal diseases have been controlled by the application of chemical fungicides. This strategy usually results in efficient control, but is also associated with the possible development of resistant pathogens and can be associated with a negative impact on the environment. Moreover, in certain crops, such as barley and wheat, the control of fungal pathogens by chemical fungicides is difficult or impractical.

Recent techniques have allowed a better understanding of the interactions between plants and their pathogens at the molecular level and mechanisms of resistance have been partially unraveled. While a large portion of this molecular characterization has been conducted in the model plant *Arabidopsis*, resistance mechanisms have also begun to be elucidated in economically important crops.

The powdery mildews are a major disease affecting most plant species and have been widely studied. They are characterized by spots or patches of a white to grayish growth on plant tissues, which correspond to the mycelium and cleistothecia of the fungus. Powdery mildews are caused by several species of fungi of the order Erysiphales. For example, *Erysiphe graminis* causes the powdery mildew of cereals and grasses. While powdery mildews are hard to control in most crops, barley lines resistant to most known isolates of the pathogen are available. It was shown that mutations at a single locus, the Mlo locus, are responsible for resistance phenotype. The mechanism of mlo resistance has been partially elucidated; it involves the formation of large cell wall appositions, called papillae, at the contact sites with the pathogen, which mainly contain callose, but also carbohydrates,

phenols and proteins. In mlo plants, cell wall appositions prevent the penetration of the pathogen, thus providing resistance.

Unfortunately, this powerful tool to control powdery mildews is restricted to barley. In the view of the problems caused by fungal diseases in agriculture, in particular by powdery mildews, there remains an unfulfilled need for new and effective strategies to control these types of pathogens in other crops, which are economically attractive to the farmers and environmentally acceptable.

The present invention addresses the need for novel disease control strategies in plants via the application of genetic engineering techniques. In particular, this invention relates to control strategies against powdery mildew, preferably in economically important crops.

The present invention relates to isolated DNA molecules encoding Mlo proteins, wherein such Mlo proteins confer resistance of plants to fungal pathogens. In particular, the invention relates to Mlo proteins containing conserved amino acid sequences that the inventors of the present invention are first to discover, and to the isolated DNA molecules encoding such Mlo proteins. The present invention is also drawn to vectors for expression of the DNA molecules of the present invention in plants. The present invention further relates to transgenic plants comprising any one of the DNA molecules of the present invention. The present invention also describes agricultural products with improved phytosanitary properties comprising transgenic plants resistant to fungal pathogens by expression of any of the DNA molecules of the present invention. The present invention also further relates to methods of making plants resistant to fungal diseases by altering the expression in transgenic plants of proteins encoded by the endogenous copies of the genes corresponding to any one of the DNA molecules of the present invention or altering the activity or the stability of proteins encoded by the endogenous copies of the genes corresponding to any one of the DNA molecules of the present invention. Such transgenic plants are desirably resistant to pathogens that infect living epidermal plant cells, more desirably to fungi from the order Erysiphales, also known as powdery mildews, preferably from the genus *Erysiphe*, the causing agent of powdery mildew, more preferably the plants are resistant to *Erysiphe graminis*. The present invention further describes a method to isolate DNA molecules encoding proteins having the same or a similar function as the DNA molecules of the present invention and encoding the conserved amino acid sequences set forth in the present invention.



The present invention thus provides new and effective strategies to control fungal diseases in economically important crops, potentially reducing amounts of chemicals applied to crops and reducing the risk of appearance of pathogens resistant to control agents.

The invention thus provides:

A DNA molecule encoding an Mlo protein conferring upon a plant resistance to fungal pathogens, wherein said protein comprises at least one amino acid sequence identical or substantially similar to an amino acid sequence set forth in SEQ ID No:1 or SEQ ID No:2, wherein said DNA molecule is preferably a cDNA molecule. In a preferred embodiment, the DNA molecule is preferably not derived from barley and is derived from a plant which is either a dicot or from a group of plants consisting of wheat, corn, rice, oats, rye, sorghum, sugarcane, millet, milo, and the palm family. In a preferred embodiment, the DNA molecule of the present invention is identical or substantially similar to any one of the nucleotide sequences set forth in SEQ ID No: 3, SEQ ID No: 5 or SEQ ID No:7, or encodes an Mlo protein identical or substantially similar to an Mlo protein set forth in SEQ ID No: 4, SEQ ID No: 6 or SEQ ID No: 8. In a more preferred embodiment, the DNA molecule comprising the nucleotide sequences set forth in SEQ ID No: 3, SEQ ID No: 5 or SEQ ID No: 7 is derived from wheat. In another preferred embodiment, the DNA molecule of the present invention is identical or substantially similar to any one of the nucleotide sequences set forth in SEQ ID No: 9, SEQ ID No: 11, SEQ ID No: 13, SEQ ID No:15 or SEQ ID No: 17, or encodes an Mlo protein identical or substantially similar to an Mlo protein encoded by any one of the nucleotide sequences set forth in SEQ ID No: 10, SEQ ID No: 12, SEQ ID No: 14, SEQ ID No:16 or SEQ ID No:18. In a more preferred embodiment, the DNA molecule comprising the nucleotide sequences set forth in SEQ ID No: 9, SEQ ID No: 11, SEQ ID No: 13, SEQ ID No:15 or SEQ ID No: 17 is derived from *Arabidopsis thaliana*. In another preferred embodiment, said DNA molecules mentioned hereinbefore are modified such that the activity of the endogenous protein is lost. In one particular embodiment of the present invention, said DNA modification results in one, all or a combination of the following changes in the amino acid sequence of the corresponding protein

- Trp (163) to Arg
- frameshift after Pro (396)
- frameshift after Trp (160)
- Met (1) to Ile

- Gly (227) to Asp
- Met (1) to Val
- Arg (11) to Trp
- missing Phe (183), Thr (184)
- Val (31) to Glu
- Ser (32) to Phe
- Leu (271) to His.

In a further preferred embodiment, the fungal pathogens desirably infect living epidermal cells, more desirably the fungal pathogens are from the order Erysiphales, also known as powdery mildews, preferably from the genus *Erysiphe* and more preferably the fungal pathogen is *Erysiphe graminis*.

In a further embodiment, the isolated DNA molecule is antisense to an isolated molecule as described above, e.g. antisense to a DNA molecule, e.g. a cDNA molecule, encoding an Mlo protein comprising at least one amino acid sequence identical or substantially similar to an amino acid sequences set forth in SEQ ID No:1 or SEQ ID No:2, especially antisense to a DNA molecule identical or substantially similar to a DNA molecule set forth in SEQ ID Nos:3, 5, 7, 9, 11, 13, 15 or 17 and encoding an Mlo protein identical or substantially similar to an Mlo protein set forth in SEQ ID Nos:4, 6, 8, 10, 12, 14, 16 or 18.

The invention further provides:

A protein comprising at least one amino acid sequence identical or substantially similar to an amino acid sequence set forth in SEQ ID No: 1 or SEQ ID No:2, wherein said protein is an Mlo protein and confers upon a plant resistance to fungal pathogens. The protein is preferably not derived from barley and is derived from a plant which is either a dicot or from a group of plants consisting of wheat, corn, rice, oats, rye, sorghum, sugarcane, millet, milo, and the palm family. In a preferred embodiment, the protein of the present invention is encoded by a nucleotide sequence identical or substantially similar to any one of the nucleotide sequences set forth in SEQ ID No: 3, SEQ ID No: 5 or SEQ ID No:7 or is identical or substantially similar to any one of the Mlo proteins set forth in SEQ ID No: 4, SEQ ID No: 6 or SEQ ID No: 8. In a more preferred embodiment, the protein is derived from wheat. In another preferred embodiment, the protein of the present invention is encoded by a nucleotide sequence identical or substantially similar to any one of the nucleotide sequences set forth in SEQ ID No: 9, SEQ ID No: 11, SEQ ID No: 13, SEQ ID No:15 or

SEQ ID No: 17 or is identical or substantially similar to any one of the Mlo proteins set forth in SEQ ID No: 10, SEQ ID No: 12, SEQ ID No: 14, SEQ ID No: 16 or SEQ ID No: 18. In a more preferred embodiment, the protein is derived from *A. thaliana*. In another preferred embodiment, the fungal pathogens desirably infect living epidermal cells, more desirably the fungal pathogens are from the order Erysiphales, also known as powdery mildews, preferably from the genus *Erysiphe* and more preferably the fungal pathogen is *Erysiphe graminis*. In a further embodiment, the present invention also encompasses mutated forms or truncated forms of proteins encoded by any of the DNA molecules described above.

The invention further provides:

An expression cassette comprising any one of the DNA molecules described above, e.g. a cDNA as described above, wherein the DNA molecule is operably linked to a promoter and to termination signals capable of expressing the DNA molecule in plants. In a preferred embodiment, the expression cassette is heterologous. In a further preferred embodiment, the promoter and the termination signals are eukaryotic. In a further preferred embodiment, the promoter and termination signals are heterologous with respect to the coding region.

The invention further provides:

A vector comprising any one of the expression cassettes described above. In a preferred embodiment, the vector is used for transformation of the expression cassette in plants. In another preferred embodiment, the vector of the present invention is used for amplification of any one of the DNA molecules described above.

The invention further provides:

A cell comprising an expression cassette or parts of it comprising an isolated DNA molecule of the present invention, wherein said DNA molecule in said expression cassette is expressible in said cell. In a preferred embodiment, the DNA molecule is not derived from barley. In another preferred embodiment, the cell is a plant cell. In a further preferred embodiment, the expression cassette is stably integrated in the genome of the cell or is included in a self-replicating vector and remains in the cell as an extrachromosomal molecule.

The invention further provides:

A plant comprising an expression cassette or parts of it comprising an isolated DNA molecule of the present invention. In a preferred embodiment, the DNA molecule is not derived from barley. In another preferred embodiment, the DNA molecule comprised in the expression cassette is expressible in the plant. In another preferred embodiment, the DNA molecule is stably integrated in the plant genome or is included in a self-replicating vector and remains in the cell as an extrachromosomal molecule. In another preferred embodiment, the plant is resistant to fungal pathogens, desirably fungal pathogens which infect living epidermal cells, more desirably the fungal pathogens are from the order Erysiphales, also known as powdery mildews, preferably from the genus *Erysiphe* and more preferably the fungal pathogen is *Erysiphe graminis*.

The inventions also relates to the seed for such a plant, which seed is optionally treated (e.g. primed or coated) and/or packaged, e.g. placed in a bag with instructions for use.

The invention further provides:

Agricultural products comprising a plant comprising an isolated DNA molecule of the present invention. In a preferred embodiment, the agricultural product is used as e.g. feed, food, or silage and does not contain mycotoxins produced by fungal pathogens, such as e.g. aflatoxins. Therefore, the agricultural product has improved phytosanitary properties.

The invention further provides:

A method for making a plant resistant to a fungal pathogen comprising the step of:

- a) expressing in a plant an RNA transcript encoded by any one of the DNA molecules described above in "sense" orientation; or
- b) expressing in a plant an RNA transcript encoded by any one of the DNA molecules described above in "anti-sense" orientation; or
- c) expressing in a plant a ribozyme capable of specifically cleaving a messenger RNA transcript encoded by an endogenous gene corresponding to any one of the DNA molecules described above; or
- d) expressing in a plant an aptamer specifically directed to an endogenous protein encoded by a gene corresponding to any one of the DNA molecules described above; or
- e) expressing in a plant a mutated or a truncated form of any one of the DNA molecules described above, so that it can act as a dominant negative mutant; or
- f) modifying by homologous recombination in a plant at least one chromosomal copy of the gene corresponding to any one of the DNA molecules described above; or

g) modifying by homologous recombination in a plant at least one chromosomal copy of the regulatory elements of a gene corresponding to any one of the DNA molecules described above.

The invention further provides:

A plant obtained by any one of the method described immediately above including the seed for such a plant, which seed is optionally treated (e.g., primed or coated) and/or packaged, e.g. placed in a bag with instructions for use. In another preferred embodiment, the plant obtained is resistant to fungal pathogens, desirably fungal pathogens which infect living epidermal cells, more desirably the fungal pathogens are from the order Erysiphales, also known as powdery mildews, preferably from the genus *Erysiphe* and more preferably the fungal pathogen is *Erysiphe graminis*.

The invention further provides:

An agricultural product with improved phytosanitary properties obtained by any one of the methods described immediately above.

The invention further provides:

A method for isolating DNA molecules encoding Mlo proteins comprising the steps of:

- a) mixing a degenerated oligonucleotide encoding at least six amino acids of SEQ ID No:1 and a degenerated oligonucleotide complementary to a sequence encoding at least six amino acids of SEQ ID No:2, with DNA extracted from a plant under conditions allowing hybridization of said degenerated oligonucleotides to said DNA; and
- b) amplifying a DNA fragment of said plant DNA, wherein said DNA fragment comprises a at its left and right ends nucleotide sequences that can anneal to said degenerated oligonucleotides in step a); and
- c) obtaining a full-length cDNA clone comprising the DNA fragment of step b).

The invention further provides:

A method for producing mutated copies of the nucleotide sequences of the present invention by "in-vitro recombination" or "DNA shuffling". The mutated copies of the nucleotide sequences of the present invention are used to confer improved resistance to fungal pathogens. In a preferred embodiment, the mutant copies of the nucleotide

sequences of the present invention are used to confer resistance to a broader range of pathogens. One such method is described below:

A method for mutagenizing a DNA molecule according to the present invention, wherein said DNA molecule has been cleaved into double-stranded-random fragments of a desired size, and comprising the steps of:

- a) adding to the resultant population of double-stranded random fragments one or more single or double-stranded oligonucleotides, wherein said oligonucleotides comprise an area of identity and an area of heterology to the double-stranded template polynucleotide;
- b) denaturing the resultant mixture of double-stranded random fragments and oligonucleotides into single-stranded fragments;
- c) incubating the resultant population of single-stranded fragments with a polymerase under conditions which result in the annealing of said single-stranded fragments at said areas of identity to form pairs of annealed fragments, said areas of identity being sufficient for one member of a pair to prime replication of the other, thereby forming a mutagenized double-stranded polynucleotide; and
- d) repeating the second and third steps for at least two further cycles, wherein the resultant mixture in the second step of a further cycle includes the mutagenized double-stranded polynucleotide from the third step of the previous cycle, and the further cycle forms a further mutagenized double-stranded polynucleotide.

## DEFINITIONS

An "isolated DNA molecule" is a nucleotide sequence that, by the hand of man, exists apart from its native environment and is therefore not a product of nature. An isolated nucleotide sequence may exist in a purified form or may exist in a non-native environment such as, for example, a transgenic host cell.

A "protein" as defined herein is the entire protein encoded by the corresponding nucleotide sequence, or is a portion of the protein encoded by the corresponding portion of the nucleotide sequence.

An "isolated protein" is a protein that is encoded by an isolated nucleotide sequence and is therefore not a product of nature. An isolated protein may exist in a purified form or may exist in a non-native environment, such as a transgenic host cell, wherein the protein would

not normally expressed or would be expressed in a different form or different amount in an isogenic non-transgenic host cell.

A plant "resistant to a fungal pathogen" has no or lesser symptoms of a fungal infection caused by the fungal , by inhibiting or limiting the ability of the fungal pathogen to grow on the plant. As a consequence, the plant grows better, has higher yields and produces more seeds.

"A protein conferring resistance upon a plant to a fungal pathogen" means that the protein is involved in the regulation of plant genetic pathways responsible for resistance of the plant to the fungal pathogen. The protein may be a positive regulator in that it enhances resistance of the plant to the fungal pathogen, or the protein may be a negative regulator in that it represses resistance of the plant to the fungal pathogen. A particular example of a protein conferring resistance upon a plant to a fungal pathogen is an Mlo protein.

An "Mlo protein" means herein a member of a family of proteins (the Mlo family) having a substantially similar function in a disease resistance pathway and sharing some structural homology. The structural homology can be e.g. that the members of the family share at least one conserved region.

In its broadest sense, the term "substantially similar", when used herein with respect to a nucleotide sequence, means a nucleotide sequence corresponding to a reference nucleotide sequence, wherein the corresponding sequence encodes a polypeptide having substantially the same structure and function as the polypeptide encoded by the reference nucleotide sequence, e.g. where only changes in amino acids not affecting the polypeptide function occur. Desirably the substantially similar nucleotide sequence encodes the polypeptide encoded by the reference nucleotide sequence. The percentage of identity between the substantially similar nucleotide sequence and the reference nucleotide sequence desirably is at least 80%, more desirably 85%, preferably at least 90%, more preferably at least 95%, still more preferably at least 99%.

The term "substantially similar", when used herein with respect to a protein, means a protein corresponding to a reference protein, wherein the protein has substantially the same structure and function as the reference protein, e.g. where only changes in amino acids not

affecting the polypeptide function occur. When used for a protein or an amino acid sequence the percentage of identity between the substantially similar and the reference protein or amino acid sequence desirably is at least 80%, more desirably 85%, preferably at least 90%, more preferably at least 95%, still more preferably at least 99%.

The percentage of sequence identity is determined using computer programs that are based on dynamic programming algorithms. Computer programs that are preferred within the scope of the present invention include the BLAST (Basic Local Alignment Search Tool) search programs designed to explore all of the available sequence databases regardless of whether the query is protein or DNA. Version BLAST 2.0 (Gapped BLAST) of this search tool has been made publicly available on the Internet (currently <http://www.ncbi.nlm.nih.gov/BLAST/>). It uses a heuristic algorithm which seeks local as opposed to global alignments and is therefore able to detect relationships among sequences which share only isolated regions. The scores assigned in a BLAST search have a well-defined statistical interpretation. Said programs are preferably run with optional parameters set to the default values.

The term "gene" refers to a coding sequence and associated regulatory sequences wherein the coding sequence is transcribed into RNA such as mRNA, rRNA, tRNA, snRNA, sense RNA or antisense RNA. Examples of regulatory sequences are promoter sequences, 5' and 3' untranslated sequences and termination sequences. Further elements that may be present are, for example, introns.

"Expression" refers to the transcription and/or translation of an endogenous gene or a transgene in plants. In the case of antisense constructs, for example, expression may refer to the transcription of the antisense DNA only.

"Expression cassette" as used herein means a DNA sequence capable of directing expression of a particular nucleotide sequence in an appropriate host cell, comprising a promoter operably linked to the nucleotide sequence of interest which is operably linked to termination signals. It also typically comprises sequences required for proper translation of the nucleotide sequence. The coding region usually codes for a protein of interest but may also code for a functional RNA of interest, for example antisense RNA or a nontranslated RNA that, in the sense or antisense direction, inhibits expression of a particular gene, e.g., antisense RNA. The expression cassette comprising the nucleotide sequence of interest



may be chimeric, meaning that at least one of its components is heterologous with respect to at least one of its other components. The expression cassette may also be one which is naturally occurring but has been obtained in a recombinant form useful for heterologous expression. Typically, however, the expression cassette is heterologous with respect to the host, i.e., the particular DNA sequence of the expression cassette does not occur naturally in the host cell and must have been introduced into the host cell or an ancestor of the host cell by a transformation event. The expression of the nucleotide sequence in the expression cassette may be under the control of a constitutive promoter or of an inducible promoter which initiates transcription only when the host cell is exposed to some particular external stimulus. In the case of a multicellular organism, such as a plant, the promoter can also be specific to a particular tissue or organ or stage of development.

"Heterologous" as used herein means "of different natural or of synthetic origin" or represent a non-natural state. For example, if a host cell is transformed with a nucleic acid sequence derived from another organism, particularly from another species, that gene is heterologous with respect to that host cell and also with respect to descendants of the host cell which carry that gene. The transforming nucleic acid may comprise a heterologous promoter, heterologous coding sequence, or heterologous termination sequence. Alternatively, the transforming nucleic acid may be completely heterologous or may comprise any possible combination of heterologous and endogenous nucleic acid sequences. Similarly, heterologous refers to a nucleotide sequence derived from and inserted into the same natural, original cell type, but which is present in a non-natural state, e.g. a different copy number, or under the control of different regulatory elements.

The term "promoter" refers to a DNA sequence that initiates transcription of an associated DNA sequence. The promoter region may also include elements that act as regulators of gene expression such as activators, enhancers, and/or repressors.

"Synthetic nucleotide sequence" as used herein means a nucleotide sequence comprising structural characters that are not present in the natural sequence. For example, an artificial sequence that resembles more closely the G+C content and the normal codon distribution of dicot and/or monocot genes is said to be synthetic.

A regulatory DNA sequence is said to be "operably linked to" or "associated with" a DNA sequence that codes for an RNA or a protein if the two sequences are situated such that the regulatory DNA sequence affects expression of the coding DNA sequence.

"Regulatory elements" refer to sequences involved in conferring the expression of a nucleotide sequence. Regulatory elements comprise a promoter operably linked to the nucleotide sequence of interest and termination signals. They also typically encompass sequences required for proper translation of the nucleotide sequence.

A "plant" refers to any plant or part of a plant and particularly to seed plants at any stage of development. Therein are also included cuttings, cell or tissue cultures and seeds. As used in conjunction with the present invention, the term "plant tissue" includes, but is not limited to, whole plants, plant organs, plant seeds, protoplasts, callus, cell cultures, and any groups of plant cells organized into structural and/or functional units.

A "plant cell" refers to the structural and physiological unit of the plant, comprising a protoplast and a cell wall. The plant cell may be in form of an isolated single cell or a cultured cell, or as a part of higher organized unit such as, for example, a plant tissue, or a plant organ.

"Transformation" as used herein means introduction of a nucleic acid into a cell. In particular, the stable integration of a DNA molecule into the genome of an organism of interest.

A "selectable marker" is conferred by a gene whose expression in a plant cell gives the cell a selective advantage. The selective advantage possessed by the cells transformed with the selectable marker gene may be due to their ability to grow in the presence of a negative selective agent, such as an antibiotic or a herbicide, compared to the growth of non-transformed cells. The selective advantage possessed by the transformed cells, compared to non-transformed cells, may also be due to their enhanced or novel capacity to utilize an added compound as a nutrient, growth factor or energy source. Selectable marker gene also refers to a gene or a combination of genes whose expression in a plant cell gives the cell both, a negative and a positive selective advantage.

A "screenable marker" is conferred by a gene whose expression does not give a selective advantage to a transformed cell, but whose expression makes the transformed cell phenotypically distinct from untransformed cells.

#### **BRIEF DESCRIPTION OF THE SEQUENCES IN THE SEQUENCE LISTING**

SEQ ID No:1	conserved amino acid sequence 1
SEQ ID No:2	conserved amino acid sequence 2
SEQ ID No:3	nucleotide sequence of wheat Mlo protein TrMlo1
SEQ ID No:4	protein sequence of TrMlo1
SEQ ID No:5	nucleotide sequence of wheat Mlo protein TrMlo2
SEQ ID No:6	protein sequence of TrMlo2
SEQ ID No:7	nucleotide sequence of wheat Mlo protein TrMlo3
SEQ ID No:8	protein sequence of TrMlo3
SEQ ID No:9	nucleotide sequence of Arabidopsis Mlo protein CIB10259
SEQ ID No:10	protein sequence of CIB10259
SEQ ID No:11	nucleotide sequence of Arabidopsis Mlo protein CIB10295
SEQ ID No:12	protein sequence of CIB10295
SEQ ID No:13	nucleotide sequence of Arabidopsis Mlo protein CIB10296
SEQ ID No:14	protein sequence of CIB10296
SEQ ID No:15	nucleotide sequence of Arabidopsis Mlo protein F19850
SEQ ID No:16	protein sequence of F19850
SEQ ID No:17	nucleotide sequence of Arabidopsis Mlo protein U95973
SEQ ID No:18	protein sequence of U95973
SEQ ID No:19	oligonucleotide MLO-1
SEQ ID No:20	oligonucleotide MLO-3
SEQ ID No:21	oligonucleotide MLO-5
SEQ ID No:22	oligonucleotide MLO-7
SEQ ID No:23	oligonucleotide MLO-10
SEQ ID No:24	oligonucleotide MLO-15
SEQ ID No:25	oligonucleotide MLO-26
SEQ ID No:26	oligonucleotide MLO-GSP1

SEQ ID No:27 oligonucleotide MLO-GSP2  
SEQ ID No:28 oligonucleotide ST27  
SEQ ID No:29 oligonucleotide N37544-1  
SEQ ID No:30 oligonucleotide N37544-2  
SEQ ID No:31 oligonucleotide T22146-1  
SEQ ID No:32 oligonucleotide T22146-2  
SEQ ID No:33 oligonucleotide H76041-1  
SEQ ID No:34 oligonucleotide H76041-2  
SEQ ID No:35 oligonucleotide SAS-1  
SEQ ID No:36 oligonucleotide SAS-2  
SEQ ID No:37 oligonucleotide SAS-3  
SEQ ID No:38 oligonucleotide SAS-4  
SEQ ID No:39 oligonucleotide SAS-5  
SEQ ID No:40 oligonucleotide SAS-6  
SEQ ID No:41 oligonucleotide SAS-7  
SEQ ID No:42 oligonucleotide SAS-8

**Deposits**

Deposited material	Accession number	Date of deposit
pCIB 10259	NRRL B-21945	3/10/98
pCIB 10295	NRRL B-21946	3/10/98
pCIB 10296	NRRL B-21947	3/10/98
TrMlo1	NRRL B-21948	3/10/98
TrMlo1-5	NRRL B-21949	3/10/98
TrMlo2	NRRL B-21950	3/10/98
TrMlo2-5	NRRL B-21951	3/10/98
TrMlo3	NRRL B-21952	3/10/98
TrMlo3-5	NRRL B-21953	3/10/98

All deposits were made with the Northern Regional Research Center, 1815 Northern University Street, Peoria, Illinois 61604, USA.

The present invention relates to DNA molecules encoding Mlo proteins, which confer upon a plant resistance to fungal pathogens. The inventors of the present invention are first to identify conserved amino acid sequences among Mlo proteins. The conserved amino acid sequences of the present invention are conserved between three Mlo proteins derived from wheat and three Mlo proteins derived from *A. thaliana*. These amino acid sequences are also conserved in two predicted Arabidopsis Mlo proteins. The first conserved amino acid sequence, which is set forth in SEQ ID No:1, comprises 13 amino acids. The fourth amino acid in SEQ ID No:1 is either L, V or I, its fifth amino acid is either V or L and its seventh amino acid is either F or L. The thirteenth amino acid in SEQ ID No:1 is not I and is preferably either T, S or A. The second conserved amino acid sequence, which is set forth in SEQ ID No:2, comprises 14 amino acids. The first amino acid in SEQ ID No:2 is not M and is preferably either I, V, S or G. Its third amino acid is either F, L or V, its sixth amino acid is either Y or N, its seventh amino acid is either A or V, its eighth amino acid is either L or I, its tenth amino acid is either T or S. The invention encompasses isolated Mlo proteins comprising at least one of the conserved amino acid sequences described above and isolated DNA molecules which encode such Mlo proteins. The invention also encompasses isolated Mlo proteins which comprise both conserved sequences set forth in SEQ ID No: 1 and SEQ ID No:2. In a preferred embodiment, the isolated DNA molecules encoding the Mlo proteins of the present invention are cDNA molecules.

In a further embodiment, the DNA molecules encoding Mlo proteins comprising at least one of the conserved amino acid sequences are not derived from barley. In another embodiment, such DNA molecules are derived from a dicot or from wheat, corn, rice, oats, rye, sorghum, sugarcane, millet, milo, or the palm family. In preferred embodiment, a DNA molecule of the present invention is identical or substantially similar to the DNA molecules set forth in SEQ ID No:3, 5 or 7 and in SEQ ID No:9, 11, 13, 15 or 17 or encode an Mlo protein, which is identical or substantially similar to any one of the Mlo proteins set forth in SEQ ID No:4, 6 or 8 or in SEQ ID No:10, 12, 14, 16, 18. The DNA molecules set forth in SEQ ID No:3, 5 or 7 are derived from wheat and encode the Mlo proteins set forth in SEQ ID No:4, 6 or 8, respectively. The isolation of such DNA molecules is further illustrated in example 1. The DNA molecules set forth in SEQ ID No:9, 11, 13, 15 or 17 are derived from Arabidopsis and encode the Mlo proteins set forth in SEQ ID No:10, 12, 14, 16 or 18, respectively. The isolation of such DNA molecules is further illustrated in example 2.

The DNA molecule of SEQ ID No:3 encoding a wheat Mlo protein called TrMlo1 is deposited as strains TrMlo1 and TrMlo1-5 with accession numbers NRRL B-21948 and NRRL B-

21949, respectively. The DNA molecule of SEQ ID No:5 encoding a wheat Mlo protein called TrMlo2 is deposited as strains TrMlo2 and TrMlo2-5 with accession numbers NRRL B-21950 and NRRL B-21951, respectively. The DNA molecule of SEQ ID No:7 encoding a wheat Mlo protein called TrMlo3 is deposited as strains TrMlo3 and TrMlo3-5 with accession numbers NRRL B-21952 and NRRL B-21953, respectively. TrMlo1 and TrMlo3 comprise the full-length cDNAs of the corresponding Mlo genes and also comprise some of the corresponding 5' and 3' untranslated regions. TrMlo2 is the longest cDNA clone of the corresponding gene that was recovered. It comprises the entire coding region with exception of the first methionin (start codon) as deduced from a comparison with TrMlo1 and TrMlo3. TrMlo2 also comprises some of the 3' untranslated region of the corresponding gene.

The DNA molecule of SEQ ID No:9 encoding an Arabidopsis Mlo protein called CIB10259 is deposited as strain pCIB10259 with accession number NRRL B-21945. The DNA molecule of SEQ ID No:11 encoding an Arabidopsis Mlo protein called CIB10295 is deposited as strain pCIB10295 with accession number NRRL B-21946. The DNA molecule of SEQ ID No:13 encoding an Arabidopsis Mlo protein called CIB10296 is deposited as strain pCIB1029 with accession number NRRL B-21947. CIB10259, CIB10295 and CIB10296 comprise the full-length cDNAs of the corresponding Mlo genes and also comprise and also comprise some of the corresponding 5' and 3' untranslated regions. The nucleotide sequences encoding the Arabidopsis Mlo protein family members F19850 and U95973 are obtained from the Genbank. However, for both clones, a predicted amino acid sequence is determined and is found not to match the amino acid sequence predicted in the Genbank entry. The Mlo proteins determined by the inventors of the present invention are therefore novel and not obvious. Both newly predicted proteins contain the conserved amino acid sequences set forth in SEQ ID No:1 and 2 and are therefore encompassed by the present invention as well as the isolated cDNAs encoding them.

An Mlo protein encoded by a DNA molecule of the present invention confer upon a plant resistance to fungal pathogens, desirably fungal pathogens which infect living epidermal plant cells, more desirably fungal pathogens from the order Erysiphales, also known as powdery mildews (Agrios G. (1988) Plant Pathology, Third Edition, Academic Press Inc., in particular p 271). Preferably, an Mlo protein encoded by a DNA molecule of the present invention confer upon a plant resistance from the genus *Erysiphe*, more preferably the fungal pathogen is *Erysiphe graminis*.

The present invention also encompasses recombinant vectors comprising any one of the DNA molecules of this invention. In these vectors, such DNA molecules are preferably comprised in an expression cassette comprising regulatory elements for expression of the DNA molecules in a host cell capable of expressing such DNA molecules. Such regulatory elements are usually a promoter and termination signals and preferably also include elements allowing efficient translation of a protein encoded by a DNA molecule of the present invention. In a preferred embodiment, an expression cassette is heterologous. Such vectors are used for transformation of the expression cassette comprising any one of the DNA molecules of this invention into a host cells. In a preferred embodiment, the expression cassette is stably integrated into the DNA of such host cell. In another preferred embodiment, the expression cassette is comprised in a vector, which is capable of replication in a host cell and remains in the host cell as an extrachromosomal molecule. In a further preferred embodiment, such extrachromosomal replicating molecule is used amplify the DNA molecules of this invention in a host cell. In a preferred embodiment, such host cell is a microorganism, such as a bacteria, in particular *E. coli*. In another preferred embodiment, a host cell is an eukaryotic cells, such as e.g. a yeast cell, an insect cell, or a plant cell.

In a further embodiment, a DNA molecule of the present invention is modified by incorporation of random mutations in a technique known as *in-vitro* recombination or DNA shuffling. This technique is described in Stemmer et al., Nature 370: 389-391 (1994) and US Patent 5,605,793 incorporated herein by reference. Millions of mutant copies of the nucleotide sequences are produced based on the original nucleotide sequence described herein and variants with improved properties, such as increased resistance for fungal pathogens or resistance against a wider range of pathogens are recovered. The method encompasses forming a mutagenized double-stranded polynucleotide from a template double-stranded polynucleotide comprising the nucleotide sequence of this invention, wherein the template double-stranded polynucleotide has been cleaved into double-stranded-random fragments of a desired size, and comprises the steps of adding to the resultant population of double-stranded random fragments one or more single or double-stranded oligonucleotides, wherein said oligonucleotides comprise an area of identity and an area of heterology to the double-stranded template polynucleotide; denaturing the resultant mixture of double-stranded random fragments and oligonucleotides into single-stranded fragments; incubating the resultant population of single-stranded fragments with a polymerase under conditions which result in the annealing of said single-stranded

fragments at said areas of identity to form pairs of annealed fragments, said areas of identity being sufficient for one member of a pair to prime replication of the other, thereby forming a mutagenized double-stranded polynucleotide; and repeating the second and third steps for at least two further cycles, wherein the resultant mixture in the second step of a further cycle includes the mutagenized double-stranded polynucleotide from the third step of the previous cycle, and the further cycle forms a further mutagenized double-stranded polynucleotide. In a preferred embodiment, the concentration of a single species of double-stranded random fragment in the population of double-stranded random fragments is less than 1% by weight of the total DNA. In a further preferred embodiment, the template double-stranded polynucleotide comprises at least about 100 species of polynucleotides. In another embodiment, the size of the double-stranded random fragments is from about 5 bp to 5 kb. In a further embodiment, the fourth step of the method comprises repeating the second and the third steps for at least 10 cycles.

The present invention also encompasses cells comprising a DNA molecule of the present invention, wherein the DNA molecule is not in its natural cellular environment. In a preferred embodiment, such cells are plant cells. In another preferred embodiment, a DNA molecule of the present invention is expressible in such cells and is comprised in an expression cassette which allow their expression in such cells. In a preferred embodiment, the expression cassette is stably integrated into the DNA of such host cell. In another preferred embodiment, the expression cassette is comprised in a vector, which is capable of replication in the cell and remains in the cell as an extrachromosomal molecule.

The present invention also encompasses a plant comprising the plant cells described above. In a further embodiment, the DNA molecules of the present invention are expressible in the plant, and expression of any one of the DNA molecules of the present invention or of a portion thereof in transgenic plants confers resistance upon the transgenic plant against fungal pathogens. In a preferred embodiment, the fungal pathogens desirably infect living epidermal cells, more desirably the fungal pathogens are from the order Erysiphales, also known as powdery mildews, preferably from the genus *Erysiphe* and more preferably the fungal pathogen is *Erysiphe graminis*. The present invention therefore also encompasses transgenic plants made resistant to fungal pathogens by the expression of any one of the DNA molecules of the present invention or of a portion thereof.



Plants transformed in accordance with the present invention may be monocots or dicots and include, but are not limited to, maize, wheat, barley, rye, sweet potato, bean, pea, chicory, lettuce, cabbage, cauliflower, broccoli, turnip, radish, spinach, asparagus, onion, garlic, pepper, celery, squash, pumpkin, hemp, zucchini, apple, pear, quince, melon, plum, cherry, peach, nectarine, apricot, strawberry, grape, raspberry, blackberry, pineapple, avocado, papaya, mango, banana, soybean, tomato, sorghum, sugarcane, sugarbeet, sunflower, rapeseed, clover, tobacco, carrot, cotton, alfalfa, rice, potato, eggplant, cucumber, *Arabidopsis thaliana*, and woody plants such as coniferous and deciduous trees, especially maize, wheat, or sugarbeet.

Once a desired nucleotide sequence has been transformed into a particular plant species, it may be propagated in that species or moved into other varieties of the same species, particularly including commercial varieties, using traditional breeding techniques.

For their expression in transgenic plants, the DNA molecules may require modification and optimization. It is known in the art that all organisms have specific preferences for codon usage, and the codons in the nucleotide sequence comprised in the DNA molecules of the present invention can be changed to conform with specific plant preferences, while maintaining the amino acids encoded thereby. Furthermore, high expression in plants is best achieved from coding sequences which have at least 35% GC content, and preferably more than 45%. Nucleotide sequences which have low GC contents may express poorly due to the existence of ATTTA motifs which may destabilize messages, and AATAAA motifs which may cause inappropriate polyadenylation. Although preferred gene sequences may be adequately expressed in both monocotyledonous and dicotyledonous plant species, sequences can be modified to account for the specific codon preferences and GC content preferences of monocotyledons or dicotyledons as these preferences have been shown to differ (Murray *et al.* Nucl. Acids Res. 17: 477-498 (1989)). In addition, the nucleotide sequences are screened for the existence of illegitimate splice sites which cause message truncation. All changes required to be made within the nucleotide sequences such as those described above are made using well known techniques of site directed mutagenesis, PCR, and synthetic gene construction using the methods described in the published patent applications EP 0 385 962, EP 0 359 472 and WO 93/07278.

For efficient initiation of translation, sequences adjacent to the initiating methionine may require modification. For example, they can be modified by the inclusion of sequences known to be effective in plants. Joshi has suggested an appropriate consensus for plants (NAR 15: 6643-6653 (1987)) and Clontech suggests a further consensus translation initiator

(1993/1994 catalog, page 210). These consensus sequences are suitable for use with the nucleotide sequences of this invention. The sequences are incorporated into constructions comprising the nucleotide sequence, up to and including the ATG (whilst leaving the second amino acid unmodified), or alternatively up to and including the GTC subsequent to the ATG (with the possibility of modifying the second amino acid of the transgene).

The DNA molecules in transgenic plants are driven by a promoter shown to be functional in plants. The choice of promoter will vary depending on the temporal and spatial requirements for expression, and also depending on the target species. For the protection of plants against foliar pathogens, expression in leaves is preferred; for the protection of plants against ear pathogens, expression in inflorescences (e.g. spikes, panicles, cobs, etc.) is preferred; for protection of plants against root pathogens, expression in roots is preferred; for protection of seedlings against soil-borne pathogens, expression in roots and/or seedlings is preferred. In many cases, however, protection against more than one type of phytopathogen is sought, and thus expression in multiple tissues is desirable. Although many promoters from dicotyledons have been shown to be operational in monocotyledons and *vice versa*, ideally dicotyledonous promoters are selected for expression in dicotyledons, and monocotyledonous promoters for expression in monocotyledons. However, there is no restriction to the provenance of selected promoters; it is sufficient that they are operational in driving the expression of the DNA molecules in the desired cell.

Preferred promoters which are expressed constitutively include promoters derived from *Agrobacterium* opine synthase genes, e.g. the nos promoter, or a dual promoter from the *Agrobacterium* Ti plasmid (Velten et al. (1984) EMBO J. 3: 2723-2730), or viral promoters operable in plants, e.g. the CaMV 35S and 19S promoters, and promoters from genes encoding actin or ubiquitin. Another preferred promoter is a synthetic promoter, such as the Gelvin Super MAS promoter (Ni et al. (1995) Plant J. 7: 661-676). The DNA molecules of this invention can also be expressed under the regulation of promoters which are chemically regulated. This enables the protein conferring fungal diseases to be synthesized only when the crop plants are treated with the inducing chemicals. Preferred technology for chemical induction of gene expression is detailed in the published application EP 0 332 104 and US patent 5,614,395. A preferred promoter for chemical induction is the tobacco PR-1a promoter.

A preferred category of promoters is that which is wound inducible. Numerous promoters have been described which are expressed at wound sites and also at the sites of

phytopathogen infection. Ideally, such a promoter should only be active locally at the sites of infection, and in this way the protein controlling fungal diseases only accumulates in cells which need to synthesize it to kill the invading insect pest. Preferred promoters of this kind include those described by Stanford *et al.* Mol. Gen. Genet. 215: 200-208 (1989), Xu *et al.* Plant Molec. Biol. 22: 573-588 (1993), Logemann *et al.* Plant Cell 1: 151-158 (1989), Rohrmeier & Lehle, Plant Molec. Biol. 22: 783-792 (1993), Firek *et al.* Plant Molec. Biol. 22: 129-142 (1993), and Warner *et al.* Plant J. 3: 191-201 (1993).

Preferred tissue specific expression patterns include green tissue specific, root specific, stem specific, and flower specific. Promoters suitable for expression in green tissue include many which regulate genes involved in photosynthesis and many of these have been cloned from both monocotyledons and dicotyledons. A preferred promoter is the maize PEPC promoter from the phosphoenol carboxylase gene (Hudspeth & Grula, Plant Molec. Biol. 12: 579-589 (1989)). A preferred promoter for root specific expression is that described by de Framond (FEBS 290: 103-106 (1991); EP 0 452 269) and a further preferred root-specific promoter is that from the T-1 gene provided by this invention. A preferred stem specific promoter is that described in US patent 5,625,136 and which drives expression of the maize *trpA* gene.

Preferred embodiments of the invention are transgenic plants expressing a DNA molecule in a root-specific fashion. Further preferred embodiments are transgenic plants expressing the DNA molecule in a wound-inducible or pathogen infection-inducible manner.

In addition to the selection of a suitable promoter, constructions for expression of the protein in plants require an appropriate transcription terminator to be attached downstream of the heterologous nucleotide sequence. Several such terminators are available and known in the art (*e.g.* *tm1* from CaMV, E9 from *rbcS*). Any available terminator known to function in plants can be used in the context of this invention.

Numerous other sequences can be incorporated into expression cassettes for the DNA molecules of this invention. These include sequences which have been shown to enhance expression such as intron sequences (*e.g.* from *Adh1* and *bronze1*) and viral leader sequences (*e.g.* from TMV, MCMV and AMV).

It may be preferable to target expression of the DNA molecules to different cellular localizations in the plant. In some cases, localization in the cytosol may be desirable, whereas in other cases, localization in some subcellular organelle may be preferred. Subcellular localization of transgene encoded enzymes can be undertaken using

techniques well known in the art. Typically, the DNA encoding the target peptide from a known organelle-targeted gene product is manipulated and fused upstream of the nucleotide sequence. Many such target sequences are known for the chloroplast and their functioning in heterologous constructions has been shown.

Vectors suitable for plant transformation are described elsewhere in this specification. For *Agrobacterium*-mediated transformation, binary vectors or vectors carrying at least one T-DNA border sequence are suitable, whereas for direct gene transfer any vector is suitable and linear DNA containing only the construction of interest may be preferred. In the case of direct gene transfer, transformation with a single DNA species or co-transformation can be used (Schocher *et al.* *Biotechnology* 4: 1093-1096 (1986)). For both direct gene transfer and *Agrobacterium*-mediated transfer, transformation is usually (but not necessarily) undertaken with a selectable marker which may provide resistance to an antibiotic (kanamycin, hygromycin or methatrexate) or a herbicide (glufosinate, glyphosate or a protoporphyrinogen oxidase inhibitor), or a selectable marker which may confer a selective advantage to transformed cells, such as a phospho-mannose isomerase gene. The choice of selectable marker is not, however, critical to the invention.

In another preferred embodiment, the DNA molecules of this invention are directly transformed into the plastid genome. Plastid transformation technology is extensively described in U.S. Patent Nos. 5,451,513, 5,545,817, and 5,545,818, in PCT application no. WO 95/16783, and in McBride *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91, 7301-7305. The basic technique for chloroplast transformation involves introducing regions of cloned plastid DNA flanking a selectable marker together with the DNA molecule of interest into a suitable target tissue, e.g., using biolistics or protoplast transformation (e.g., calcium chloride or PEG mediated transformation). The 1 to 1.5 kb flanking regions, termed targeting sequences, facilitate homologous recombination with the plastid genome and thus allow the replacement or modification of specific regions of the plastome. Initially, point mutations in the chloroplast 16S rRNA and rps12 genes conferring resistance to spectinomycin and/or streptomycin are utilized as selectable markers for transformation (Svab, Z., Hajdukiewicz, P., and Maliga, P. (1990) *Proc. Natl. Acad. Sci. USA* 87, 8526-8530; Staub, J. M., and Maliga, P. (1992) *Plant Cell* 4, 39-45). This resulted in stable homoplasmic transformants at a frequency of approximately one per 100 bombardments of target leaves. The presence of cloning sites between these markers allowed creation of a plastid targeting vector for introduction of foreign genes (Staub, J.M., and Maliga, P. (1993) *EMBO J.* 12, 601-606,

hereby incorporated by reference). Substantial increases in transformation frequency are obtained by replacement of the recessive rRNA or r-protein antibiotic resistance genes with a dominant selectable marker, the bacterial *aadA* gene encoding the spectinomycin-detoxifying enzyme aminoglycoside-3'-adenyltransferase (Svab, Z., and Maliga, P. (1993) *Proc. Natl. Acad. Sci. USA* 90, 913-917). Previously, this marker had been used successfully for high-frequency transformation of the plastid genome of the green alga *Chlamydomonas reinhardtii* (Goldschmidt-Clermont, M. (1991) *Nucl. Acids Res.* 19, 4083-4089). Other selectable markers useful for plastid transformation are known in the art and encompassed within the scope of the invention. Typically, approximately 15-20 cell division cycles following transformation are required to reach a homoplastidic state. Plastid expression, in which genes are inserted by homologous recombination into all of the several thousand copies of the circular plastid genome present in each plant cell, takes advantage of the enormous copy number advantage over nuclear-expressed genes to permit expression levels that can readily exceed 10% of the total soluble plant protein.

The present invention also encompasses agricultural products, which comprise transgenic plants made resistant to fungal pathogens by expression of any one of the DNA molecules of the present invention or made resistant to fungal pathogens by any one of the methods described *infra*. Since such plants are resistant to fungal pathogens, pathogen growth on them is suppressed. Such plants and the agricultural products derived thereof are therefore less likely to contain mycotoxins, which are naturally produced by many fungal pathogens and which can be very toxic to human and to animals. Therefore, such agricultural products have better phytosanitary properties. In a preferred embodiment, such agricultural products are used as feedstuff, as silage, or as food.

It is a further object of the present invention, to provide methods of making plants resistant to fungal pathogens. The Mlo proteins encoded by the DNA molecules of the present invention confer upon a plant resistance to a fungal pathogen and it is a preferred object of the present invention to alter the expression of such proteins in their natural host environment. It is a further preferred object of the present invention to alter the stability or the activity of such proteins in their natural environment. Such alterations of the expression, stability, or activity of the proteins encoded by the DNA molecules of the present invention in a plant result in increased resistance of the plant to fungal pathogens. In a preferred embodiment, a protein encoded by a DNA molecule of the present invention the protein is a

negative regulator of resistance of a plant to fungal pathogens, in that it represses genetic pathways in the plant which are responsible for resistance of the plant to fungal pathogens. Therefore, it is a preferred object of the present invention to reduce the expression of the Mlo proteins encoded by the DNA molecules of the present invention in their natural host environment, or to reduce the stability or the activity of such proteins in their natural host environment.

#### "Sense" Suppression

In a preferred embodiment, reduction of the expression of a protein encoded by a DNA molecule of the present invention is obtained by "sense" suppression (referenced in e.g. Jorgensen et al. (1996) Plant Mol. Biol. 31, 957-973). In this case, the entirety or a portion of a DNA molecule of the present invention is comprised in an expression cassette, which is introduced in a host cell, preferably a plant cell, in which the DNA molecule is expressible. The DNA molecule is inserted in the expression cassette in the "sense orientation", meaning that the 5' end of the DNA molecule is adjacent to the promoter in the expression cassette and that the coding strand of the DNA molecule can be transcribed. In a preferred embodiment, the DNA molecule is fully translatable and all the genetic information comprised in the DNA molecule or a portion thereof is translated into a protein. In another preferred embodiment, the DNA molecule is partially translatable and a short peptide is translated. In a preferred embodiment, this is achieved by inserting at least one premature stop codon in the DNA molecule, which brings translation to a halt. In another more preferred embodiment, the DNA molecule is transcribed but no translation product is being made. This is usually achieved by removing the start codon, e.g. the "ATG", of the protein encoded by the DNA molecule. In a further preferred embodiment, the expression cassette comprising the DNA molecule or a portion thereof is stably integrated in the genome of the host cell. In another preferred embodiment, the expression cassette comprising the DNA molecule or a portion thereof is comprised in an extrachromosomal replicating molecule. In transgenic plants containing one of the expression cassettes described immediately above, the expression of the gene corresponding to the DNA molecule comprised in the expression cassette is reduced or abolished, leading to reduced levels of the protein or to its absence in the transgenic plants. As a result thereof, the transgenic plants are resistant to fungal pathogens.

### **"Anti-sense" Suppression**

In another preferred embodiment, the reduction of the expression of a protein encoded by a DNA molecule of the present invention is obtained by "anti-sense" suppression. The entirety or a portion of a DNA molecule of the present invention is comprised in an expression cassette, wherein the DNA molecule is introduced in a host cell, preferably a plant cell, in which the DNA molecule is expressible. The DNA molecule is inserted in the expression cassette in the "anti-sense orientation", meaning that the 3' end of the DNA molecule is adjacent to the promoter in the expression cassette and that the non-coding strand of the DNA molecule can be transcribed. In a preferred embodiment, the expression cassette comprising the DNA molecule or a portion thereof is stably integrated in the genome of the host cell. In another preferred embodiment, the expression cassette comprising the DNA molecule or a portion thereof is comprised in an extrachromosomal replicating molecule. Several publications describing this approach are cited for further illustration (Green, P. J. et al., *Ann. Rev. Biochem.* 55:569-597 (1986); van der Krol, A. R. et al, *Antisense Nuc. Acids & Proteins*, pp. 125-141 (1991); Abel, P. P. et al., *Proc. Natl. Acad. Sci. USA* 86:6949-6952 (1989); Ecker, J. R. et al., *Proc. Natl. Acad. Sci. USA* 83:5372-5376 (Aug. 1986)).

### **Homologous Recombination**

In another preferred embodiment, at least one genomic copy corresponding to a DNA molecule of the present invention is modified in the genome of the plant by homologous recombination as further illustrated in Paszkowski et al., *EMBO Journal* 7:4021-26 (1988). This technique uses the property of homologous sequences to recognize each other and to exchange nucleotide sequences between each by a process known in the art as homologous recombination. Homologous recombination can occur between the chromosomal copy of a nucleotide sequence in a cell and an incoming copy of the nucleotide sequence introduced in the cell by transformation. Specific modifications are thus accurately introduced in the chromosomal copy of the nucleotide sequence. In one embodiment, the regulatory elements of the gene encoding a protein of the present invention are modified. The existing regulatory elements are replaced by different regulatory elements, thus reducing expression of the protein, or they are mutated or deleted, thus abolishing the expression of the protein. In another embodiment, the coding region of the protein is modified by deletion of a part of the coding sequence of the entire coding sequence, or by mutation. Expression of a mutated protein can also confer upon the plant increased resistance to fungal pathogens.

In another preferred embodiment, a mutation in the chromosomal copy of a DNA molecule is introduced by transforming a cell with a chimeric oligonucleotide composed of a contiguous stretch of RNA and DNA residues in a duplex conformation with double hairpin caps on the ends. An additional feature of the oligonucleotide is the presence of 2'-O-methylation at the RNA residues. The RNA/DNA sequence is designed to align with the sequence of a chromosomal copy of a DNA molecule of the present invention and to contain the desired nucleotide change. This technique is further illustrated in US patent 5,501,967.

#### **Ribozymes**

In a further embodiment, the RNA coding for a protein of the present invention is cleaved by a catalytic RNA, or ribozyme, specific for such RNA. The ribozyme is expressed in transgenic plants and results in reduced amounts of RNA coding for the protein of the present invention in plant cells, thus leading to reduced amounts of protein accumulated in the cells and increased resistance of the plant to fungal pathogens. This method is further illustrated in US patent 4,987,071.

#### **Dominant-Negative Mutants**

In another preferred embodiment, the activity of the proteins encoded by the nucleotide sequences of this invention is changed. This is achieved by expression of dominant negative mutants of the proteins in transgenic plants, leading to the loss of activity of the endogenous protein.

#### **Aptamers**

In a further embodiment, the activity of a protein encoded by a DNA molecule of the present invention is inhibited by expressing in transgenic plants nucleic acid ligands, so-called aptamers, which specifically bind to the protein. Aptamers are preferentially obtained by the SELEX (Systematic Evolution of Ligands by EXponential Enrichment) method. In the SELEX method, a candidate mixture of single stranded nucleic acids having regions of randomized sequence is contacted with the protein and those nucleic acids having an increased affinity to the target are partitioned from the remainder of the candidate mixture. The partitioned nucleic acids are amplified to yield a ligand enriched mixture. After several iterations a nucleic acid with optimal affinity to the protein is obtained and is used for expression in transgenic plants. This method is further illustrated in US patent 5,270,163.



**Methods for Isolating Nucleotide Sequences Comprising the Conserved Sequences**

The conserved sequences comprised in Mlo proteins encoded by the DNA molecules of the present invention are used for the isolation of other DNA molecules encoding such sequences. In a preferred embodiment, a mixture of degenerated oligonucleotides that contain at least one possible oligonucleotide encoding a sequences set forth in SEQ ID No:1, or in SEQ ID No:2 is produced. The mixture of oligonucleotides encoding the sequences set forth in SEQ ID No:1 and the mixture of oligonucleotides complementary to sequences encoding the sequences set forth in SEQ ID No:2 are used in a PCR amplification reaction with a template DNA of choice. Mixtures of degenerated oligonucleotides are well known in the art and the degree of degeneration is varied as needed. In a preferred embodiment, the template DNA is a sample of total DNA from a plant, wherein such DNA sample is obtained by methods well known in the art. The amplified fragments resulting from the PCR reactions described above are isolated by methods well known in the art and are used to isolate the corresponding full-length cDNAs by screening a cDNA library or by using a RACE protocol, both well known in the art. This method represents a novel and useful strategy to isolate new genes conferring upon plants resistance to fungal pathogens.

The invention will be further described by reference to the following detailed examples. These examples are provided for purposes of illustration only, and are not intended to be limiting unless otherwise specified.

**EXAMPLES****Example 1: Cloning and Sequencing of Mlo Genes from Wheat**

Mlo genes from wheat are cloned using a reverse transcription-PCR approach. RNA is prepared from leaves of wheat cultivar UC703 and used to program reverse transcription using a Stratagene RT-PCR kit. The resulting cDNA is employed in PCR reactions using the following primers:

MLO-26 5' TTC CAG CAC CGG CAC AAG AA 3' (SEQ ID No:25)

MLO-10 5' AAG AAC TGC CTG AAG AAG GC 3' (SEQ ID No:23)  
MLO-7 5' CAG AAA CTT GTC TCA TCC CTG G 3' (SEQ ID No:22)  
MLO-5 5' ACA GAG ACC ACC TCC TTG GAA 3' (SEQ ID No:21)  
MLO-15 5' CAC CAC CTT CAT GAT GCT CA 3' (SEQ ID No:24)

PCR is performed using the primer pairs listed below, the reactions of which resulted in amplification of fragments of the indicated sizes:

MLO-26 and MLO-10	503 bp
MLO-26 and MLO-7	1481 bp
MLO-5 and MLO-15	650 bp

The fragments are cloned into either pCR2.1 or pCR2.1-TOPO (Invitrogen). Plasmid DNA is prepared from the transformants and subjected to DNA sequencing. Sequencing reveals the existence of three different cDNA sequences with very high similarity to each other. These wheat Mlo genes are called TrMlo1, TrMlo2 and TrMlo3.

An additional wheat Mlo clone is isolated by screening a wheat cDNA library constructed in Lambda-ZAPII vector. To screen this library, mass excision is performed to convert the library into a Bluescript-based collection of cDNA clones. These are grown separately in pools of 80,000 independent clones, and plasmid DNA prepared. PCR reactions are performed on the pooled DNAs using oligonucleotide primers MLO-5 and MLO-15 (see above). Three pools yield a band of the expected size of 650 base pairs. These pools are subsequently fractionated by serial culturing of bacterial clones at lower and lower density, followed at each step by plasmid DNA preparation and PCR on each subpool using primers MLO-5 and MLO-15. After several rounds of fractionation a single clone is isolated which contained an insert with an Mlo sequence. Sequencing the inserts of two of these clones reveals that they contain identical inserts. The sequence of the third clone reveals an insert identical to that of the other two clones, but with 40 additional bases at the 5' end.

Cloning of the remaining portions of the wheat Mlo is accomplished by random amplification of cDNA ends (RACE). RACE reactions are performed on wheat UC703 poly-A<sup>+</sup> RNA using a Marathon cDNA amplification kit (Clontech). Poly-A<sup>+</sup> RNA is prepared from total

wheat RNA using an oligo-dT cellulose column (Gibco BRL). Oligonucleotides used for the RACE reactions are:

MLO-GSP1      5' TGG ACC TCT TCA TGT TCG ATC CCA TCT G 3' (SEQ ID No:26)

MLO-GSP2      5' CCT GAC GCT GTT CCA GAA TGC GTT TCA 3' (SEQ ID No:27)

Amplification using primer MLO-GSP1 and the 5' adapter primer provided in the kit results in a DNA fragment of ~1300 nucleotides. Amplification using primer MLO-GSP2 and the 3' adapter results in a DNA fragment of ~600 nucleotides. The fragments are cloned into pCR2.1-TOPO and called TrMlo1-5, TrMlo2-5 and TrMlo3-5 and comprise the 5' end of the wheat Mlo genes TrMlo1, TrMlo2 and TrMlo3, respectively. Plasmid DNA is prepared from the clones and the plasmid inserts are sequenced.

#### **Example 2: Cloning of Mlo cDNAs from Arabidopsis**

Comparison of the Mlo protein sequence to translations of database entries using the program TBLASTN revealed a number of entries with similarity to Mlo. Of these, full-length cDNAs corresponding to three Arabidopsis EST entries are cloned, with accession numbers H76041, N37544, and T22146. For each EST, oligonucleotides are designed to amplify a sequence corresponding to the EST from an Arabidopsis cDNA library constructed in plasmid pFL61 (Minet et al. (1992) Gene Nov 16; 121(2): 393-396.). The oligonucleotides used are:

N37544-1 5' AAG ATC AAG ATG AGG ACG TGG AAG TCG TGG 3' (SEQ ID No:29)

N37544-2 5' AGG CTG AAC CAC TGG GGC GCC TCT CAC CAC 3' (SEQ ID No:30)

T22146-1 5' CAA GTA TAT GAT GCG CGC TCT AGA GGA TGA 3' (SEQ ID No:31)

T22146-2 5' AGG TTT CAC CAC TAA GTC TCC TTC AAT GGC 3' (SEQ ID No:32)

H76041-1 5' GAT CAT TCA AGA CTT AGG CTC ACT CAT GAG 3' (SEQ ID No:33)

H76041-2 5' AAC AGC AAG GAA GAT TAC AAA TGA TGC CCA 3' (SEQ ID No:34)

The primers N37544-1 and N37544-2 amplify a ~500 base pair fragment from DNA prepared from the cDNA library, and the primers T22146-1 and T22146-2 amplify a ~250 base pair fragment. Primers H76041-1 and H76041-2 amplify two fragments of ~350 and a ~300 base pairs. The ~300 base fragment is of the size predicted from the EST sequence, and is used subsequently to diagnose the presence of a cDNA corresponding to the

H76041 EST in the cDNA library. DNA from the library is transformed into *E. coli*, and clones organized into pools of ~20,000 clones each. DNA from individual pools are screened by PCR using the different primer pairs, and positive pools subsequently subfractionated into smaller and smaller clone numbers. Isolation of individual positive clones is accomplished either by carrying this process through to completion, or in some cases by colony hybridization using EST sequences as probes once pool sizes has reached 200 clones or fewer. For ESTs N37544 and T22146, a clone corresponding to the EST is successfully isolated, and the insert sequenced. The cDNA corresponding to EST N37544 is designated CIB10295 in plasmid pCIB10295 and the plasmid containing the cDNA corresponding to EST T22146 (CIB10296) is designated pCIB10296. For these two ESTs, corresponding genomic sequences are available in the form of portions of Arabidopsis BAC clone sequences recently deposited to GenBank. Notably, the protein sequence predicted to be translated from these genomic sequences, as determined by GenBank, does not correspond to the sequence determined by directly sequencing of the cDNAs. Therefore, the amino acid sequence of the genes corresponding to the ESTs N37544 and T22146 are not obvious from the GenBank entry and are only elucidated through cloning and sequencing of the cDNA clones. The clone isolated using primers for the H76041 EST is found to contain not the gene for H76041, but a novel Mlo gene family member as insert. The insert is completely sequenced, and this Mlo gene family member is designated CIB10259 in plasmid pCIB10259.

### **Example 3: Construction of Vectors for Expression of the Mlo Genes in Wheat**

Two vectors are constructed for "antisense" expression of barley Mlo gene in wheat. PCR is performed using barley cDNA and the primer pair MLO-5 and MLO-7 (see (1) above), and the reaction results in amplification of a 1124 bp fragment which is cloned into pGEM-T (Promega). This fragment is excised from pGEM-T using the enzymes SacII and NotI. The 1124 base pair fragment is cloned into pBluescript-SK(+). The insert is now excised with BamHI and SacI restriction sites and cloned into BamHI-SacI-digested pCIB9806 (described in patent application 08/838,219) in an orientation in which the Mlo coding sequence runs opposite to the maize ubiquitin promoter. This plasmid is designated pCK01.

To construct a vector for "antisense" expression of the entire Mlo gene in wheat, PCR is performed using the primer pair MLO-1 (5' ATG TCG GAC AAA AAA GGG GT 3' (SEQ ID No:19)) and MLO-10 (see (1) above), and the reaction results in amplification of a 635 bp fragment which is cloned into pCR2.1 (Invitrogen). This fragment is excised from pCR2.1

as an EcoRI fragment, and inserted into pGEM-9Zf(-) (Promega). A 320 nucleotide fragment spanning the naturally-occurring SacI site in Mlo to primer site MLO-10 is excised with SacI and BstXI. pCK01 is digested with SacI and BstXI, and the 320 base fragment inserted. To complete construction of the Mlo gene in the monocot expression vector, a 210 nucleotide SacI fragment is excised from the pGEM-9Zf(-) derivative. This fragment contains the 5' end of the Mlo coding sequence, from the primer site MLO-1 to the naturally-occurring SacI site in the Mlo gene. The pCK01 derivative is digested with SacI and the 210 base fragment inserted. Clones are analyzed for orientation of the 210 base fragment in the newly-constructed vector by PCR, using primers MLO-1 and MLO-10. Only clones in which the 210 base fragment is inserted in the antisense orientation relative to the ubiquitin promoter yielded a 530 base pair product corresponding to the 5' end of the Mlo coding sequence. The resulting plasmid contained the entire Mlo coding sequence in "antisense" orientation relative to the ubiquitin promoter, and is designated pCK02.

To construct a vector for expression of the Mlo gene in "sense" orientation, plasmid pCK02 is digested with BamHI to release the Mlo coding sequence as insert. The BamHI fragment is religated back into the pCK02 base vector. Colonies with the Mlo coding sequence in reverse orientation relative to pCK02 are identified by SacI digestion, which yields a 1.8 kb fragment in such clones, as opposed to a 210 base fragment in clones of identical configuration as pCK02. A clone with the Mlo coding sequence in "sense" orientation relative to the maize ubiquitin promoter is selected as designated pCK03.

#### **Example 4: Construction of Vectors for Expression of the Mlo Genes in Arabidopsis**

Mlo clones in pCIB10259, pCIB10295 and pCIB10296, along with pCK02 (for the barley Mlo gene) are used in PCR reactions, resulting in bands that carry the full-length gene sequences flanked by BamHI restriction sites. The sequences of the primers used are:

SAS-1: 5' GGA TTA AGA TCT AAT GGC 3' (SEQ ID No:35, for pCIB10295)

SAS-2: 5' CAA AGA TCT TCA TTT CTT AAA AG 3' (SEQ ID No:36, for pCIB10295)

SAS-3: 5' GCG GAT CCA TGT CGG ACA AAA AAG G 3' (SEQ ID No:37, for barley Mlo)

SAS-4: 5' GCG GAT CCT CAT CCC TGG CTG AAG G 3' (SEQ ID No:38, for barley Mlo)

SAS-5: 5' GGA TCC ACC ATG GCC ACA AGA TG 3' (SEQ ID No:39, for pCIB10259)

SAS-6: 5' GGA TCC TTA GTC AAT ATC ATT AGC 3' (SEQ ID No:40, for pCIB10259)

SAS-7: 5' GCG GAT CCA TGG GTC ACG GAG GAG AAG 3' (SEQ ID No:41, for pCIB10269)

SAS-8: 5' GCG GAT CCT CAG TTG TTA TGA TCA GGA 3' (SEQ ID No:42, for pCIB10296)

The bands are cloned into pCR2.1-TOPO, and the inserts sequenced from the resulting plasmids to confirm the absence of mutations introduced by PCR. The plasmids are digested with BamHI, and the inserts purified and cloned into BamHI-digested pPEH28, a shuttle vector containing a copy of the Arabidopsis ubiquitin gene promoter UBQ3 (Norris et al. (1993) Plant Molecular Biology 21: 895-906) immediately downstream of the BamHI site. Clones containing Mlo sequences fused to UBQ3 are identified, and restriction analysis performed to identify clones with the inserts in the "sense" and "antisense" orientations relative to UBQ3. For each Mlo gene, a clone with the insert in "sense" orientation and a clone with the insert in "antisense" orientation is digested with XbaI, and the insert purified and cloned into XbaI-digested pCIB200. This places the UBQ3-Mlo gene fusion between T-DNA borders.

#### **Example 5: Transformation of Wheat and Identification of Expressors**

Wheat is transformed by particle bombardment of immature embryos as described in detail in patent application WO 94/13822. Plantlets are regenerated on media containing Basta and subjected to PCR analysis. For diagnosis of the presence of the Mlo transgenes by PCR, the following primers are used:

MLO-3: 5' ATG CTA CCA CAC GCA GAT CG 3'  
ST27: 5' ACT TCT GCA GGT CGA CTC TA 3'

The primer MLO-3 corresponds to a region of the Mlo transgene, while primer ST27 lies within the maize ubiquitin promoter sequence. The use of both Mlo gene and ubiquitin promoter primers in PCR eliminates false positives arising from the use of two Mlo primers, which might prime from the chromosomal copy of the Mlo gene present in wheat.

Plants confirmed to contain Mlo transgenes are subjected to RNA gel blot analysis, to determine whether they contain altered levels of chromosomally-encoded wheat Mlo mRNAs. Poly-A<sup>+</sup> RNAs are prepared from individual transgenic lines, and blotted onto Hybond-N<sup>+</sup> filters. Blots are probed with a 530 base fragment corresponding to the 5' end of the Mlo gene. This region is absent from the pCK01 clone; therefore, hybridization to antisense RNA expressed from the transgene in transgenic lines containing pCK01 does not occur. For pCK02 transgenic lines, in which the transgene contains this 5' end

fragment, the probe hybridizes to two bands of differing size. A ~2.5 kb mRNA corresponding to the antisense transgene is distinguished from the 2.0 kb mRNA derived from wheat chromosomally-encoded Mlo genes. The abundance of the 2.0 kb mRNA is monitored as a measurement of the efficiency of gene suppression achieved by the transgene in individual lines.

#### **Example 6: Disease Testing of Transgenic Wheat Lines**

Plants of transgenic and untransformed UC703 (control) wheat lines are grown in the greenhouse until they are two weeks old. Plants are moved to a Percival growth chamber (cycle of 8 hours dark, 16° C, and 16 hours light, 20° C), and inoculated with *Erysiphe graminis f. sp. tritici* by liberal application of spores. The degree of fungal sporulation is scored two weeks after inoculation. Plants are scored 1 (little to no hyphal growth and no visible sporulation), 2 (some hyphal growth and sporulation, but less control plants) or 3 (hyphal growth and sporulation comparable to controls). Transgenic wheat lines expressing Mlo-constructs show increased resistance to the pathogen. An example for results obtained with the antisense barley Mlo construct is shown below.

#### Screening of siblings from transgenic lines R1 and R2 for disease resistance

Siblings of Mlo antisense transformants (T2 seeds) are planted, inoculated with *E. graminis*, and scored for disease resistance.

	<b>R1</b>	<b>R2</b>	<b>UC703 (control)</b>
<b><u>Seed Planted</u></b>	24	24	24
<b>Germinated</b>	14	24	21
<b>Disease Score 1</b>	4	2	0
<b>Disease Score 2</b>	0	0	0
<b>Disease Score 3</b>	10	22	21

The fact that a small percentage of the R1 and R2 plants exhibit resistance may be due to the fact that T2 populations were tested which are still segregating for the transgene.

#### **Example 7: Analysis of Arabidopsis Lines Expressing the Mlo Genes**

Derivatives of pCIB200 containing Mlo genes are used to transform *Arabidopsis* ecotype Ws-O by vacuum infiltration (Bechtold, N., Ellis, J. and Pelletier, G. (1993) *C. R. Acad. Sci. Paris* 316, 1194-1199). Progeny are screened by kanamycin selection to identify transformants. For Mlo transgenic lines, plants are identified which expressed Mlo by RNA gel blot analysis. For Mlo genes, transformants are analyzed for alteration in the steady-state level of mRNA accumulation by RNA gel blot analysis. Transformants exhibiting sense or antisense suppression of target genes are assayed for alteration in reaction to the phytopathogenic fungi *Erysiphe cichoracearum* and *Peronospora parasitica*, and the bacterial pathogen *Pseudomonas syringae* pv. *tomato*. Leaves of transgenic plants are inspected both macroscopically and microscopically using trypan blue staining, to test for the presence of necrosis.

For *Erysiphe* inoculation, spores are liberally applied to *Arabidopsis* rosettes and plants are maintained in a Percival growth chamber at 25°C. Degree of fungal sporulation is scored 10 days after inoculation.

#### **Example 8: Use of Regions of Similarity between Mlo Sequences to Isolate Additional Mlo Gene Family Members**

Alignment of the amino acid sequences predicted to be encoded by Mlo genes revealed a number of short regions of high amino similarity between all the gene products. Degenerate primers are designed to these regions, and PCR reactions with these primers are carried out according to PCR reagent supplier's recommendations. The amplified fragments are used as a probe to isolate full-length cDNA or genomic clones of the novel Mlo genes. Amino acid sequences conserved between the Mlo proteins of the present invention (in bold) and degenerated oligonucleotides used for the isolation of additional Mlo genes are shown below:

	<b>E</b>	<b>L</b>	<b>M</b>	<b>X1</b>	<b>X2</b>	<b>G</b>	<b>X3</b>	<b>I</b>	<b>S</b>	<b>L</b>	<b>L</b>	<b>L</b>	<b>X4</b>
WHEAT													
TrMlo1	GAG	CTC	ATG	CTG	GTG	GGC	TTC	ATC					
TrMlo2	GAG	CTG	ATG	CTG	GTG	GGG	TTC	ATC					
TrMlo3	GAG	CTG	ATG	CTG	GTG	GGA	TTC	ATC					
ARABIDOPSIS													
CIB10259	GAG	CTG	ATG	ATT	CTA	GGA	TTC	ATT					



CIB10295 GAG CTT ATG CTG TTG GGA TTC ATA  
 CIB10296 GAG CTG ATG TTG TTA GGG TTT ATA  
 F19850 GAG CTG ATG GTT CTT GGA TTC ATC  
 U95973 GAG TTG ATG TTG CTG GGA CTT ATA

5' GAG CTB ATG MTB BTR GGM TTC AT 3'

	X5	T	X6	P	L	X7	X8	X9	V	X10	Q	M	G	S
<b>WHEAT</b>														
TrMlo1							GCG	CTC	GTC	ACA	CAG	ATG	GGA	TCA
TrMlo2							GCG	CTC	GTC	ACA	CAG	ATG	GGA	TCG
TrMlo3							GCG	CTA	GTC	ACA	CAG	ATG	GGA	TCA
<b>ARABIDOPSIS</b>														
CIB10259							GCA	CTA	GTT	ACT	CAG	ATG	GGT	TCA
CIB10295							GCA	CTT	GTT	ACT	CAG	ATG	GGT	AGT
CIB10296							GCC	ATC	GTC	TCA	CAG	ATG	GGA	AGT
F19850							GCA	CTC	GTA	ACT	CAG	ATG	GGT	TCT
U95973							GTA	ATC	GTT	ACT	CAG	ATG	GGA	TCT

5' WCC CAT CTG AGT GAC DAG BGC RTA 3'

X1= L, V or I, X2= V or L, X3= F or L, X4= T, S or A.

X5= I, V, S or G, X6= F, L or V, X7= Y or N, X8= A or V, X9= L or I, X10= T or S.

R=A,G Y=C,T M=A,C K=G,T S=C,G W=A,T H=A,C,T B=C,G,T V=A,C,G D=A,G,T

N=A,C,G,T.

#### Example 9: Modification of Coding Sequences and Adjacent Sequences

The DNA molecules described in this application can be modified for expression in transgenic plant hosts to achieve and optimize or down-regulate their expression. The following problems may be encountered and the modification of these DNA molecules can be undertaken using techniques well known in the art.

- (1) Codon Usage. The preferred codon usage in some plants differs from the preferred codon usage in some other plant species. Typically plant evolution has tended towards a strong preference of the nucleotides C and G in the third base position of monocotyledons,

whereas dicotyledons often use the nucleotides A or T at this position. By modifying a gene to incorporate preferred codon usage for a particular target transgenic species, many of the problems described below for GC/AT content and illegitimate splicing will be overcome.

(2) GC/AT Content. Plant genes typically have a GC content of more than 35%. DNA molecules which are rich in A and T nucleotides can cause several problems in plants. Firstly, motifs of ATTTA are believed to cause destabilization of messages and are found at the 3' end of many short-lived mRNAs. Secondly, the occurrence of polyadenylation signals such as AATAAA at inappropriate positions within the message is believed to cause premature truncation of transcription. In addition, monocotyledons may recognize AT-rich sequences as splice sites (see below).

(3) Sequences Adjacent to the Initiating Methionine. It is believed that ribosomes attach to the 5' end of the message and scan for the first available ATG at which to start translation. Nevertheless, it is believed that there is a preference for certain nucleotides adjacent to the ATG and that expression of DNA molecules of the present invention can be enhanced by the inclusion of a new consensus translation initiator at the ATG. Clontech (1993/1994 catalog, page 210) have suggested a sequence as a consensus translation initiator for the expression of the *E. coli uidA* gene in plants. Further, Joshi (NAR 15: 6643-6653 (1987)) has compared many plant sequences adjacent to the ATG and suggests a consensus sequence. In situations where difficulties are encountered in the expression of DNA molecules in plants, inclusion of one of these sequences at the initiating ATG may improve translation. In such cases the last three nucleotides of the consensus may not be appropriate for inclusion in the modified sequence due to their modification of the second AA residue. Preferred sequences adjacent to the initiating methionine may differ between different plant species. A survey of 14 maize genes located in the GenBank database provided the following results:

Position Before the Initiating ATG in 14 Maize Genes:

	<u>-10</u>	<u>-9</u>	<u>-8</u>	<u>-7</u>	<u>-6</u>	<u>-5</u>	<u>-4</u>	<u>-3</u>	<u>-2</u>	<u>-1</u>
<b>C</b>	3	8	4	6	2	5	6	0	10	7
<b>T</b>	3	0	3	4	3	2	1	1	1	0
<b>A</b>	2	3	1	4	3	2	3	7	2	3

G 6 3 6 0 6 5 4 6 1 5

This analysis can be done for the desired plant species into which the nucleotide sequence is being incorporated, and the sequence adjacent to the ATG modified to incorporate the preferred nucleotides.

(4) Removal of Illegitimate Splice Sites. DNA molecules of the present invention may also contain motifs which may be recognized in plants as 5' or 3' splice sites, and be cleaved, thus generating truncated or deleted messages. These sites can be removed using the techniques well known in the art.

(5) Creation of Dominant-Negative Mutants

In addition, DNA molecules of the present invention may also include molecules that are modified in such a way that the activity of the proteins encoded by the nucleotide sequences of this invention is changed. This is achieved by expression of dominant negative mutants of the proteins in transgenic plants, leading to the loss of activity of the endogenous protein. The location of mutations in the Mlo nucleotide sequence leading to the production of such dominant-negative mutations are listed below. Either a single mutation or a combination of the different mutations listed below can be introduced.

Mutation Number	Description of mutation	Location of mutation in wheat Mlo proteins of SEQ ID Nos: 4, 6 and 8 (aa)	Location of mutation in wheat Mlo genes (nt number)		
			TrMlo1 SEQ ID No:3	TrMlo2 SEQ ID No:5	TrMlo3 SEQ ID No:7
1	T to A	Trp (163) to Arg	662	487	684
2	deletion	frameshift after Pro (396)	1366-1367	1191-1192	1388-1389
3	deletion	frameshift after Trp (160)	656-666	481-491	678-688
4	G to A	Met (1) to Ile	178	3	200
5	G to A	Gly (227) to Asp	855	680	877

6	A to G	Met (1) to Val	176	1	198
7	A to T	Arg (11) to Trp	206	31	228
8	deletion	missing Phe (183), Thr (184)	721-726	546-551	743-748
9	T to A	Val (31) to Glu	267	92	289
10	C to T	Ser (32) to Phe	270	95	292
11	T to A	Leu (271) to His	987	812	1009

Techniques for the modification of coding sequences and adjacent sequences are well known in the art. In cases where the initial expression of a DNA molecule of the present invention is low and it is deemed appropriate to make alterations to the sequence as described above, then the construction of synthetic genes can be accomplished according to methods well known in the art. These are, for example, described in the published patent disclosures EP 0 385 962, EP 0 359 472 and WO 93/07278. In most cases it is preferable to assay the expression of gene constructions using transient assay protocols (which are well known in the art) prior to their transfer to transgenic plants.

#### **Example 10: Construction of Plant Transformation Vectors**

Numerous transformation vectors are available for plant transformation, and the DNA molecules of this invention can be used in conjunction with any such vectors. The selection of vector for use will depend upon the preferred transformation technique and the target species for transformation. For certain target species, different antibiotic or herbicide selection markers may be preferred. Selection markers used routinely in transformation include the *npfI* gene which confers resistance to kanamycin, paromomycin, geneticin and related antibiotics (Vieira and Messing, 1982, Gene 19: 259-268; Bevan et al., 1983, Nature 304:184-187) the bacterial *aadA* gene (Goldschmidt-Clermont, 1991, Nucl. Acids Res. 19: 4083-4089), encoding aminoglycoside 3'-adenylyltransferase and conferring resistance to streptomycin or spectinomycin, the *hph* gene which confers resistance to the antibiotic hygromycin (Blochlinger and Diggelmann, 1984, Mol. Cell. Biol. 4: 2929-2931), and the *dhfr* gene, which confers resistance to methotrexate (Bourouis and Jarry, 1983, EMBO J. 2: 1099-1104). Other markers to be used include a phosphinothricin acetyltransferase gene, which confers resistance to the herbicide phosphinothricin (White et al., 1990, Nucl. Acids Res. 18: 1062; Spencer et al. 1990, Theor. Appl. Genet. 79: 625-631), a mutant EPSP

synthase gene encoding glyphosate resistance (Hinchee et al., 1988, Bio/Technology 6: 915-922), a mutant acetolactate synthase (ALS) gene which confers imidazolinone or sulfonylurea resistance (Lee et al., 1988, EMBO J. 7: 1241-1248), a mutant psbA gene conferring resistance to atrazine (Smeda et al., 1993, Plant Physiol. 103: 911-917), or a mutant protoporphyrinogen oxidase gene as described in US Patent No. 5,767, 373.

Selection markers resulting in positive selection, such as a phosphomannose isomerase gene, as described in US Patent No 5,767,378, are also used.

Identification of transformed cells may also be accomplished through expression of screenable marker genes such as genes coding for chloramphenicol acetyl transferase (CAT),  $\beta$ -glucuronidase (GUS), luciferase, and green fluorescent protein (GFP) or any other protein that confers a phenotypically distinct trait to the transformed cell.

#### (1) Construction of Vectors Suitable for *Agrobacterium* Transformation

Many vectors are available for transformation using *Agrobacterium tumefaciens*. These typically carry at least one T-DNA border sequence and include vectors such as pBIN19 (Bevan, Nucl. Acids Res. (1984)) and pXYZ. Below the construction of two typical vectors is described.

##### Construction of pCIB200 and pCIB2001

The binary vectors pCIB200 and pCIB2001 are used for the construction of recombinant vectors for use with *Agrobacterium* and is constructed in the following manner. pTJS75kan is created by *NarI* digestion of pTJS75 (Schmidhauser & Helinski, J Bacteriol. 164: 446-455 (1985)) allowing excision of the tetracycline-resistance gene, followed by insertion of an *AccI* fragment from pUC4K carrying an NPTII (Vieira & Messing, Gene 19: 259-268 (1982); Bevan et al., Nature 304: 184-187 (1983); McBride et al., Plant Molecular Biology 14: 266-276 (1990)). *XhoI* linkers are ligated to the *EcoRV* fragment of pCIB7 which contains the left and right T-DNA borders, a plant selectable *nos/nptII* chimeric gene and the pUC polylinker (Rothstein et al., Gene 53: 153-161 (1987)), and the *XhoI*-digested fragment is cloned into *Sall*-digested pTJS75kan to create pCIB200 (see also EP 0 332 104, example 19). pCIB200 contains the following unique polylinker restriction sites: *EcoRI*, *SstI*, *KpnI*, *BglII*, *XbaI*, and *Sall*. pCIB2001 is a derivative of pCIB200 which created by the insertion into the polylinker of additional restriction sites. Unique restriction sites in the polylinker of pCIB2001 are *EcoRI*, *SstI*, *KpnI*, *BglII*, *XbaI*, *Sall*, *MluI*, *BclI*, *AvrII*, *Apal*, *HpaI*, and *StuI*. pCIB2001, in addition to containing these unique restriction sites also has plant and

bacterial kanamycin selection, left and right T-DNA borders for *Agrobacterium*-mediated transformation, the RK2-derived *trfA* function for mobilization between *E. coli* and other hosts, and the *OriT* and *OriV* functions also from RK2. The pCIB2001 polylinker is suitable for the cloning of plant expression cassettes containing their own regulatory signals.

#### Construction of pCIB10 and Hygromycin Selection Derivatives thereof

The binary vector pCIB10 contains a gene encoding kanamycin resistance for selection in plants, T-DNA right and left border sequences and incorporates sequences from the wide host-range plasmid pRK252 allowing it to replicate in both *E. coli* and *Agrobacterium*. Its construction is described by Rothstein *et al.* (Gene 53: 153-161 (1987)). Various derivatives of pCIB10 have been constructed which incorporate the gene for hygromycin B phosphotransferase described by Gritz *et al.* (Gene 25: 179-188 (1983)). These derivatives enable selection of transgenic plant cells on hygromycin only (pCIB743), or hygromycin and kanamycin (pCIB715, pCIB717).

#### (2) Construction of Vectors Suitable for non-*Agrobacterium* Transformation.

Transformation without the use of *Agrobacterium tumefaciens* circumvents the requirement for T-DNA sequences in the chosen transformation vector and consequently vectors lacking these sequences can be utilized in addition to vectors such as the ones described above which contain T-DNA sequences. Transformation techniques which do not rely on *Agrobacterium* include transformation via particle bombardment, protoplast uptake (*e.g.* PEG and electroporation) and microinjection. The choice of vector depends largely on the preferred selection for the species being transformed. Below, the construction of some typical vectors is described.

#### Construction of pCIB3064

pCIB3064 is a pUC-derived vector suitable for direct gene transfer techniques in combination with selection by the herbicide Basta (or phosphinothricin). The plasmid pCIB246 comprises the CaMV 35S promoter in operational fusion to the *E. coli* GUS gene and the CaMV 35S transcriptional terminator and is described in the PCT published application WO 93/07278. The 35S promoter of this vector contains two ATG sequences 5' of the start site. These sites are mutated using standard PCR techniques in such a way as to remove the ATGs and generate the restriction sites *SspI* and *PvuII*. The new restriction sites are 96 and 37 bp away from the unique *Sall* site and 101 and 42 bp away from the actual start site. The resultant derivative of pCIB246 is designated pCIB3025. The GUS

gene is then excised from pCIB3025 by digestion with *Sall* and *SacI*, the termini rendered blunt and religated to generate plasmid pCIB3060. The plasmid pJIT82 is obtained from the John Innes Centre, Norwich and the a 400 bp *SmaI* fragment containing the *bar* gene from *Streptomyces viridochromogenes* is excised and inserted into the *HpaI* site of pCIB3060 (Thompson *et al.* EMBO J 6: 2519-2523 (1987)). This generated pCIB3064 which comprises the *bar* gene under the control of the CaMV 35S promoter and terminator for herbicide selection, a gene for ampicillin resistance (for selection in *E. coli*) and a polylinker with the unique sites *SphI*, *PstI*, *HindIII*, and *BamHI*. This vector is suitable for the cloning of plant expression cassettes containing their own regulatory signals.

#### Construction of pSOG19 and pSOG35

pSOG35 is a transformation vector which utilizes the *E. coli* gene dihydrofolate reductase (DHFR) as a selectable marker conferring resistance to methotrexate. PCR is used to amplify the 35S promoter (~800 bp), intron 6 from the maize Adh1 gene (~550 bp) and 18 bp of the GUS untranslated leader sequence from pSOG10. A 250 bp fragment encoding the *E. coli* dihydrofolate reductase type II gene is also amplified by PCR and these two PCR fragments are assembled with a *SacI-PstI* fragment from pBI221 (Clontech) which comprised the pUC19 vector backbone and the nopaline synthase terminator. Assembly of these fragments generated pSOG19 which contains the 35S promoter in fusion with the Intron 6 sequence, the GUS leader, the DHFR gene and the nopaline synthase terminator. Replacement of the GUS leader in pSOG19 with the leader sequence from Maize Chlorotic Mottle Virus (MCMV) generated the vector pSOG35. pSOG19 and pSOG35 carry the pUC gene for ampicillin resistance and have *HindIII*, *SphI*, *PstI* and *EcoRI* sites available for the cloning of foreign sequences.

#### **Example 11: Requirements for Construction of Plant Expression Cassettes**

Gene sequences intended for expression in transgenic plants are firstly assembled in expression cassettes behind a suitable promoter and upstream of a suitable transcription terminator.

#### Promoter Selection

The selection of promoter used in expression cassettes will determine the spatial and temporal expression pattern of the transgene in the transgenic plant. Selected promoters will express transgenes in specific cell types (such as leaf epidermal cells, mesophyll cells, root cortex cells) or in specific tissues or organs (roots, leaves or flowers, for example) and

this selection will reflect the desired location of biosynthesis of a DNA molecule of the present invention. Alternatively, the selected promoter may drive expression of the gene under a light-induced or other temporally regulated promoter. A further alternative is that the selected promoter be chemically regulated. This provides the possibility of inducing the expression of the nucleotide sequence only when desired and caused by treatment with a chemical inducer.

#### Transcriptional Terminators

A variety of transcriptional terminators are available for use in expression cassettes. These are responsible for the termination of transcription beyond the transgene and its correct polyadenylation. Appropriate transcriptional terminators and those which are known to function in plants and include the CaMV 35S terminator, the *tm1* terminator, the nopaline synthase terminator, the pea *rbcS* E9 terminator. These can be used in both monocotyledons and dicotyledons.

#### Sequences for the Enhancement or Regulation of Expression

Numerous sequences have been found to enhance gene expression from within the transcriptional unit and these sequences can be used in conjunction with the genes of this invention to increase their expression in transgenic plants.

Various intron sequences have been shown to enhance expression, particularly in monocotyledonous cells. For example, the introns of the maize *Adh1* gene have been found to significantly enhance the expression of the wild-type gene under its cognate promoter when introduced into maize cells. Intron 1 is found to be particularly effective and enhanced expression in fusion constructs with the chloramphenicol acetyltransferase gene (Callis *et al.*, *Genes Develop* **1**: 1183-1200 (1987)). In the same experimental system, the intron from the maize *bronze1* gene had a similar effect in enhancing expression (Callis *et al.*, *supra*). Intron sequences have been routinely incorporated into plant transformation vectors, typically within the non-translated leader.

A number of non-translated leader sequences derived from viruses are also known to enhance expression, and these are particularly effective in dicotyledonous cells. Specifically, leader sequences from Tobacco Mosaic Virus (TMV, the ' $\Omega$ -sequence'), Maize Chlorotic Mottle Virus (MCMV), and Alfalfa Mosaic Virus (AMV) have been shown to be effective in enhancing expression (*e.g.* Gallie *et al.* *Nucl. Acids Res.* **15**: 8693-8711 (1987); Skuzeski *et al.* *Plant Molec. Biol.* **15**: 65-79 (1990))

#### Targeting of the Gene Product Within the Cell



Various mechanisms for targeting gene products are known to exist in plants and the sequences controlling the functioning of these mechanisms have been characterized in some detail. For example, the targeting of gene products to the chloroplast is controlled by a signal sequence found at the aminoterminal end of various proteins and which is cleaved during chloroplast import yielding the mature protein (*e.g.* Comai *et al.* J. Biol. Chem. 263: 15104-15109 (1988)). These signal sequences can be fused to heterologous gene products to effect the import of heterologous products into the chloroplast (van den Broeck *et al.* Nature 313: 358-363 (1985)). DNA encoding for appropriate signal sequences can be isolated from the 5' end of the cDNAs encoding the RUBISCO protein, the CAB protein, the EPSP synthase enzyme, the GS2 protein and many other proteins which are known to be chloroplast localized.

Other gene products are localized to other organelles such as the mitochondrion and the peroxisome (*e.g.* Unger *et al.* Plant Molec. Biol. 13: 411-418 (1989)). The cDNAs encoding these products can also be manipulated to effect the targeting of heterologous gene products to these organelles. Examples of such sequences are the nuclear-encoded ATPases and specific aspartate amino transferase isoforms for mitochondria. Targeting to cellular protein bodies has been described by Rogers *et al.* (Proc. Natl. Acad. Sci. USA 82: 6512-6516 (1985)).

In addition sequences have been characterized which cause the targeting of gene products to other cell compartments. Aminoterminal sequences are responsible for targeting to the ER, the apoplast, and extracellular secretion from aleurone cells (Koehler & Ho, Plant Cell 2: 769-783 (1990)). Additionally, aminoterminal sequences in conjunction with carboxyterminal sequences are responsible for vacuolar targeting of gene products (Shinshi *et al.* Plant Molec. Biol. 14: 357-368 (1990)).

By the fusion of the appropriate targeting sequences described above to transgene sequences of interest it is possible to direct the transgene product to any organelle or cell compartment. For chloroplast targeting, for example, the chloroplast signal sequence from the RUBISCO gene, the CAB gene, the EPSP synthase gene, or the GS2 gene is fused in frame to the aminoterminal ATG of the transgene. The signal sequence selected should include the known cleavage site and the fusion constructed should take into account any amino acids after the cleavage site which are required for cleavage. In some cases this requirement may be fulfilled by the addition of a small number of amino acids between the cleavage site and the transgene ATG or alternatively replacement of some amino acids within the transgene sequence. Fusions constructed for chloroplast import can be tested

for efficacy of chloroplast uptake by *in vitro* translation of *in vitro* transcribed constructions followed by *in vitro* chloroplast uptake using techniques described by (Bartlett *et al.* In: Edelman *et al.* (Eds.) Methods in Chloroplast Molecular Biology, Elsevier. pp 1081-1091 (1982); Wasmann *et al.* Mol. Gen. Genet. 205: 446-453 (1986)). These construction techniques are well known in the art and are equally applicable to mitochondria and peroxisomes. The choice of targeting which may be required for the insecticidal toxin. This will usually be cytosolic or chloroplastic, although it may in some cases be mitochondrial or peroxisomal. The expression of the nucleotide sequence may also require targeting to the ER, the apoplast or the vacuole.

The above described mechanisms for cellular targeting can be utilized not only in conjunction with their cognate promoters, but also in conjunction with heterologous promoters so as to effect a specific cell targeting goal under the transcriptional regulation of a promoter which has an expression pattern different to that of the promoter from which the targeting signal derives.

#### **Example 12: Examples of Expression Cassette Construction**

The present invention encompasses the expression of a DNA molecule under the regulation of any promoter which is expressible in plants, regardless of the origin of the promoter.

Furthermore, the invention encompasses the use of any plant-expressible promoter in conjunction with any further sequences required or selected for the expression of the DNA molecule. Such sequences include, but are not restricted to, transcriptional terminators, extraneous sequences to enhance expression (such as introns [*e.g.* *Adh* intron 1], viral sequences [*e.g.* TMV- $\Omega$ ]), and sequences intended for the targeting of the gene product to specific organelles and cell compartments.

##### **Constitutive Expression: the CaMV 35S Promoter**

Construction of the plasmid pCGN1761 is described in the published patent application EP 0 392 225. pCGN1761 contains the 'double' 35S promoter and the *tm1* transcriptional terminator with a unique *EcoRI* site between the promoter and the terminator and has a pUC-type backbone. A derivative of pCGN1761 is constructed which has a modified polylinker which includes *NotI* and *XhoI* sites in addition to the existing *EcoRI* site. This derivative is designated pCGN1761ENX. pCGN1761ENX is useful for the cloning of cDNA sequences or gene sequences (including microbial ORF sequences) within its polylinker for the purposes of their expression under the control of the 35S promoter in transgenic plants.

The entire 35S promoter-gene sequence-*tml* terminator cassette of such a construction can be excised by *HindIII*, *SphI*, *Sall*, and *XbaI* sites 5' to the promoter and *XbaI*, *BamHI* and *BglI* sites 3' to the terminator for transfer to transformation vectors such as those described above in example 35. Furthermore, the double 35S promoter fragment can be removed by 5' excision with *HindIII*, *SphI*, *Sall*, *XbaI*, or *PstI*, and 3' excision with any of the polylinker restriction sites (*EcoRI*, *NotI* or *XhoI*) for replacement with another promoter.

#### Modification of pCGN1761ENX by Optimization of the Translational Initiation Site

For any of the constructions described in this section, modifications around the cloning sites can be made by the introduction of sequences which may enhance translation. This is particularly useful when genes derived from microorganisms are to be introduced into plant expression cassettes as these genes may not contain sequences adjacent to their initiating methionine which may be suitable for the initiation of translation in plants. In cases where genes derived from microorganisms are to be cloned into plant expression cassettes at their ATG it may be useful to modify the site of their insertion to optimize their expression. Modification of pCGN1761ENX is described by way of example to incorporate one of several optimized sequences for plant expression (e.g. Joshi, *supra*).

#### Expression under a Chemically Regulatable Promoter

This section describes the replacement of the double 35S promoter in pCGN1761ENX with any promoter of choice; by way of example the chemically regulated PR-1a promoter is described. The promoter of choice is preferably excised from its source by restriction enzymes, but can alternatively be PCR-amplified using primers which carry appropriate terminal restriction sites. Should PCR-amplification be undertaken, then the promoter should be resequenced to check for amplification errors after the cloning of the amplified promoter in the target vector. The chemically regulatable tobacco PR-1a promoter is cleaved from plasmid pCIB1004 (see EP 0 332 104, example 21 for construction) and transferred to plasmid pCGN1761ENX. pCIB1004 is cleaved with *NcoI* and the resultant 3' overhang of the linearized fragment is rendered blunt by treatment with T4 DNA polymerase. The fragment is then cleaved with *HindIII* and the resultant PR-1a promoter containing fragment is gel purified and cloned into pCGN1761ENX from which the double 35S promoter has been removed. This is done by cleavage with *XhoI* and blunting with T4 polymerase, followed by cleavage with *HindIII* and isolation of the larger vector-terminator containing fragment into which the pCIB1004 promoter fragment is cloned. This generates a pCGN1761ENX derivative with the PR-1a promoter and the *tml* terminator and an

intervening polylinker with unique *EcoRI* and *NotI* sites. The DNA molecule of the present invention can be inserted into this vector, and the fusion products (*i.e.* promoter-gene-terminator) can subsequently be transferred to any selected transformation vector, including those described in this application.

Constitutive Expression: the Actin Promoter

Several isoforms of actin are known to be expressed in most cell types and consequently the actin promoter is a good choice for a constitutive promoter. In particular, the promoter from the rice *Act1* gene has been cloned and characterized (McElroy *et al.* Plant Cell 2: 163-171 (1990)). A 1.3 kb fragment of the promoter is found to contain all the regulatory elements required for expression in rice protoplasts. Furthermore, numerous expression vectors based on the *Act1* promoter have been constructed specifically for use in monocotyledons (McElroy *et al.* Mol. Gen. Genet. 231: 150-160 (1991)). These incorporate the *Act1*-intron 1, *Adh1* 5' flanking sequence and *Adh1*-intron 1 (from the maize alcohol dehydrogenase gene) and sequence from the CaMV 35S promoter. Vectors showing highest expression are fusions of 35S and the *Act1* intron or the *Act1* 5' flanking sequence and the *Act1* intron. Optimization of sequences around the initiating ATG (of the GUS reporter gene) also enhanced expression. The promoter expression cassettes described by McElroy *et al.* (Mol. Gen. Genet. 231: 150-160 (1991)) can be easily modified for the expression of the DNA molecules of the present invention and are particularly suitable for use in monocotyledonous hosts. For example, promoter containing fragments can be removed from the McElroy constructions and used to replace the double 35S promoter in pCGN1761ENX, which is then available for the insertion of specific gene sequences. The fusion genes thus constructed can then be transferred to appropriate transformation vectors. In a separate report the rice *Act1* promoter with its first intron has also been found to direct high expression in cultured barley cells (Chibbar *et al.* Plant Cell Rep. 12: 506-509 (1993)).

Constitutive Expression: the Ubiquitin Promoter

Ubiquitin is another gene product known to accumulate in many cell types and its promoter has been cloned from several species for use in transgenic plants (*e.g.* sunflower - Binet *et al.* Plant Science 79: 87-94 (1991), maize - Christensen *et al.* Plant Molec. Biol. 12: 619-632 (1989)). The maize ubiquitin promoter has been developed in transgenic monocot systems and its sequence and vectors constructed for monocot transformation are disclosed in the patent publication EP 0 342 926. Further, Taylor *et al.* (Plant Cell Rep. 12: 491-495 (1993))

describe a vector (pAHC25) which comprises the maize ubiquitin promoter and first intron and its high activity in cell suspensions of numerous monocotyledons when introduced via microprojectile bombardment. The ubiquitin promoter is clearly suitable for the expression of a DNA molecule of the present invention in transgenic plants, especially monocotyledons. Suitable vectors are derivatives of pAHC25 or any of the transformation vectors described in this application, modified by the introduction of the appropriate ubiquitin promoter and/or intron sequences.

#### Root Specific Expression

A preferred pattern of expression for the nucleotide sequence of the instant invention is root expression. Root expression is particularly useful for the control of soil-borne fungal pathogens. A suitable root promoter is that described by de Framond (FEBS 290: 103-106 (1991)) and also in the published patent application EP 0 452 269. This promoter is transferred to a suitable vector such as pCGN1761ENX for the insertion of the nucleotide sequence and subsequent transfer of the entire promoter-gene-terminator cassette to a transformation vector of interest.

#### Wound Inducible Promoters

Wound-inducible promoters are particularly suitable for the expression of a DNA molecules of the present invention because they are typically active not just on wound induction, but also at the sites of phytopathogen infection. Numerous such promoters have been described (*e.g.* Xu *et al.* Plant Molec. Biol. 22: 573-588 (1993), Logemann *et al.* Plant Cell 1: 151-158 (1989), Rohrmeier & Lehle, Plant Molec. Biol. 22: 783-792 (1993), Firek *et al.* Plant Molec. Biol. 22: 129-142 (1993), Warner *et al.* Plant J. 3: 191-201 (1993)) and all are suitable for use with the instant invention. Logemann *et al.* (*supra*) describe the 5' upstream sequences of the dicotyledonous potato *wun1* gene. Xu *et al.* (*supra*) show that a wound inducible promoter from the dicotyledon potato (*pin2*) is active in the monocotyledon rice. Further, Rohrmeier & Lehle (*supra*) describe the cloning of the maize *Wip1* cDNA which is wound induced and which can be used to isolated the cognate promoter using standard techniques. Similarly, Firek *et al.* (*supra*) and Warner *et al.* (*supra*) have described a wound induced gene from the monocotyledon *Asparagus officinalis* which is expressed at local wound and pathogen invasion sites. Using cloning techniques well known in the art, these promoters can be transferred to suitable vectors, fused to the DNA molecule of this invention, and used to express these genes at the sites of phytopathogen infection.

#### Pith Preferred Expression

Patent Application WO 93/07278 (to Ciba-Geigy) describes the isolation of the maize *trpA* gene which is preferentially expressed in pith cells. The gene sequence and promoter extend up to -1726 from the start of transcription are presented. Using standard molecular biological techniques, this promoter or parts thereof, can be transferred to a vector such as pCGN1761 where it can replace the 35S promoter and be used to drive the expression of a DNA molecule of this invention in a pith-preferred manner. In fact fragments containing the pith-preferred promoter or parts thereof can be transferred to any vector and modified for utility in transgenic plants.

#### Pollen-Specific Expression

Patent Application WO 93/07278 (to Ciba-Geigy) further describes the isolation of the maize calcium-dependent protein kinase (CDPK) gene which is expressed in pollen cells. The gene sequence and promoter extend up to 1400 bp from the start of transcription. Using standard molecular biological techniques, this promoter or parts thereof, can be transferred to a vector such as pCGN1761 where it can replace the 35S promoter and be used to drive the expression of a DNA molecule of this invention. In fact fragments containing the pollen-specific promoter or parts thereof can be transferred to any vector and modified for utility in transgenic plants.

#### Leaf-Specific Expression

A maize gene encoding phosphoenol carboxylase (PEPC) has been described by Hudspeth & Grula (Plant Molec Biol 12: 579-589 (1989)). Using standard molecular biological techniques the promoter for this gene can be used to drive the expression of any gene in a leaf-specific manner in transgenic plants.

#### Expression with Chloroplast Targeting

Chen & Jagendorf (J. Biol. Chem. 268: 2363-2367 (1993) have described the successful use of a chloroplast transit peptide for import of a heterologous transgene. This peptide used is the transit peptide from the *rbcS* gene from *Nicotiana plumbaginifolia* (Poulsen *et al.* Mol. Gen. Genet. 205: 193-200 (1986)). Using the restriction enzymes *DraI* and *SphI*, or *Tsp509I* and *SphI* the DNA sequence encoding this transit peptide can be excised from plasmid prbcS-8B (Poulsen *et al. supra*) and manipulated for use with any of the constructions described above. The *DraI-SphI* fragment extends from -58 relative to the initiating *rbcS* ATG to, and including, the first amino acid (also a methionine) of the mature peptide immediately after the import cleavage site, whereas the *Tsp509I-SphI* fragment

extends from -8 relative to the initiating *rbcS* ATG to, and including, the first amino acid of the mature peptide. Thus, these fragment can be appropriately inserted into the polylinker of any chosen expression cassette generating a transcriptional fusion to the untranslated leader of the chosen promoter (*e.g.* 35S, PR-1a, actin, ubiquitin *etc.*), whilst enabling the insertion of the DNA molecule of this invention in correct fusion downstream of the transit peptide. Constructions of this kind are routine in the art. For example, whereas the *DraI* end is already blunt, the 5' *Tsp509I* site may be rendered blunt by T4 polymerase treatment, or may alternatively be ligated to a linker or adaptor sequence to facilitate its fusion to the chosen promoter. The 3' *SphI* site may be maintained as such, or may alternatively be ligated to adaptor or linker sequences to facilitate its insertion into the chosen vector in such a way as to make available appropriate restriction sites for the subsequent insertion of the DNA molecule of this invention. Ideally the ATG of the *SphI* site is maintained and comprises the first ATG of the DNA molecule of this invention. Chen & Jagendorf (*supra*) provide consensus sequences for ideal cleavage for chloroplast import, and in each case a methionine is preferred at the first position of the mature protein. At subsequent positions there is more variation and the amino acid may not be so critical. In any case, fusion constructions can be assessed for efficiency of import *in vitro* using the methods described by Bartlett *et al.* (In: Edelman *et al.* (Eds.) *Methods in Chloroplast Molecular Biology*, Elsevier. pp 1081-1091 (1982)) and Wasmann *et al.* (*Mol. Gen. Genet.* **205**: 446-453 (1986)). Typically the best approach may be to generate fusions using the DNA molecule of this invention with no modifications at the aminotermminus, and only to incorporate modifications when it is apparent that such fusions are not chloroplast imported at high efficiency, in which case modifications may be made in accordance with the established literature (Chen & Jagendorf, *supra*; Wasman *et al.*, *supra*; Ko & Ko, *J. Biol. Chem.* **267**: 13910-13916 (1992)).

Similar manipulations can be undertaken to utilize other GS2 chloroplast transit peptide encoding sequences from other sources (monocotyledonous and dicotyledonous) and from other genes. In addition, similar procedures can be followed to achieve targeting to other subcellular compartments such as mitochondria.

***What is claimed is:***

1. A DNA molecule encoding an Mlo protein, wherein said Mlo protein comprises at least one of the sequences set forth in SEQ ID No:1 or SEQ ID No:2 and wherein said Mlo protein confers upon a plant resistance to a fungal pathogen.
2. The DNA molecule of claim 1, wherein said DNA molecule is identical or substantially similar to any one of the nucleotide sequences set forth in SEQ ID No:3, SEQ ID No:5 or SEQ ID No:7, or encodes an Mlo protein identical or substantially similar to an Mlo protein set forth in SEQ ID No:4, SEQ ID No:6 or SEQ ID No:8.
3. The DNA molecule of claim 1, wherein said DNA molecule is identical or substantially similar to any one of the nucleotide sequences set forth in SEQ ID No:9, SEQ ID No:11, SEQ ID No:13, SEQ ID No:15 or SEQ ID No:17, or encodes an Mlo protein identical or substantially similar to an Mlo protein encoded set forth in SEQ ID No:10, SEQ ID No:12, SEQ ID No:14, SEQ ID No:16 or SEQ ID No:18.
4. The DNA molecule of any one of claims 1 to 3, wherein said DNA molecule is not derived from barley.
5. The DNA molecule of any one of claims 1 to 4, wherein said DNA is modified such that the activity of the endogenous protein is lost.
6. The DNA molecule of claim 5, wherein the DNA modification results in one, all or a combination of the following changes in the amino acid sequence of the corresponding protein
  - Trp (163) to Arg
  - frameshift after Pro (396)
  - frameshift after Trp (160)
  - Met (1) to Ile
  - Gly (227) to Asp
  - Met (1) to Val
  - Arg (11) to Trp



- missing Phe (183), Thr (184)
- Val (31) to Glu
- Ser (32) to Phe
- Leu (271) to His

7. A DNA molecule antisense to a DNA molecule of any one of claims 1 to 6.

8. A protein comprising at least one of the sequences set forth in SEQ ID No:1 or SEQ ID No:2, wherein said protein is an Mlo protein and confers upon a plant resistance to a fungal pathogen.

9. The protein of claim 8, wherein said protein is encoded by a nucleotide sequence identical or substantially similar to any one of the sequences set forth in SEQ ID No:3, SEQ ID No:5 or SEQ ID No:7, or is identical or substantially similar to any one of the proteins set forth in SEQ ID No:4, SEQ ID No:6 or SEQ ID No:8.

10. The protein of claim 8, wherein said protein is encoded by a nucleotide sequence identical or substantially similar to any one of the sequences set forth in SEQ ID No:9, SEQ ID No:11, SEQ ID No:13, SEQ ID No:15 or SEQ ID No:17, or is identical or substantially similar to any one of the proteins set forth in SEQ ID No:10, SEQ ID No:12, SEQ ID No:14, SEQ ID No:16 or SEQ ID No:18.

11. The protein of any one of claims 8 to 10, wherein said protein is not derived from barley.

12. An expression cassette comprising a DNA molecule of any one of claims 1 to 7.

13. A vector comprising an expression cassette comprising a DNA molecule of claim 12.

14. A cell comprising an expression cassette or parts of it comprising a DNA molecule of any one of claims 1 to 7, wherein said DNA molecule in said expression cassette is expressible in said cell.

15. The cell of claim 14, wherein said DNA molecule is stably integrated in the genome of said cell.

16. The cell of any one of claims 14 or 15, wherein said cell is a plant cell.

17. A plant comprising an expression cassette or parts of it comprising a DNA molecule of any one of claims 1 to 7, wherein said DNA molecule in said expression cassette is expressible in said plant.

18. The plant of claim 17, wherein said DNA molecule is stably integrated in the plant genome of said plant.

19. An agricultural product comprising a plant comprising an isolated DNA molecule of any one of claims 1 to 7, wherein said agricultural product has improved phytosanitary properties.

20. A method for making a plant resistant to a fungal pathogen comprising the steps of:

- a) expressing in said plant a DNA molecule of any one of claims 1 to 6 in "sense" orientation; or
- b) expressing in said plant a DNA molecule of any one of claims 1 to 6 in "anti-sense" orientation; or
- c) expressing in said plant a ribozyme capable of specifically cleaving a messenger RNA transcript encoded by an endogenous gene corresponding to a DNA molecule of any one of claims 1 to 6; or
- d) expressing in a plant an aptamer specifically directed to a protein or part of a protein encoded by a DNA molecule of any one of claims 1 to 6; or
- e) expressing in a plant a mutated or a truncated form of a DNA molecule of any one of claims 1 to 6; or
- f) modifying by homologous recombination in a plant at least one chromosomal copy of the gene corresponding to a DNA molecule of any one of claims 1 to 6.

21. A plant made resistant to a fungal pathogen by the method of claim 20.

22. The plant of claim 21, wherein said fungal pathogen infects living epidermal cells.
23. The plant of claim 21, wherein said fungal pathogen is from the order Erysiphales.
24. The plant of claim 21, wherein said fungal pathogen is from the genus *Erysiphe*.
25. The plant of claim 21, wherein said fungal pathogen is *Erysiphe graminis*.
26. An agricultural product with improved phytosanitary properties obtained using the method of claim 20.
27. A method for isolating DNA molecules encoding Mlo proteins comprising the steps of:
- a) mixing a degenerated oligonucleotide encoding at least six amino acids of SEQ ID No:1 and a degenerated oligonucleotide complementary to a sequence encoding at least six amino acids of SEQ ID No:2, with DNA extracted from a plant under conditions allowing hybridization of said degenerated oligonucleotides to said DNA; and
  - b) amplifying a DNA fragment of said plant DNA, wherein said DNA fragment comprises at its left and right ends nucleotide sequences that can anneal to said degenerated oligonucleotides in step a); and
  - c) obtaining a full-length cDNA clone comprising the DNA fragment of step b).
28. A method for mutagenizing a DNA molecule of claim 1, wherein said DNA molecule has been cleaved into double-stranded-random fragments of a desired size, and comprising the steps of:
- a) adding to the resultant population of double-stranded random fragments one or more single or double-stranded oligonucleotides, wherein said oligonucleotides comprise an area of identity and an area of heterology to the double-stranded template polynucleotide;
  - b) denaturing the resultant mixture of double-stranded random fragments and oligonucleotides into single-stranded fragments;
  - c) incubating the resultant population of single-stranded fragments with a polymerase under conditions which result in the annealing of said single-stranded fragments at said areas of identity to form pairs of annealed fragments, said areas of identity being

sufficient for one member of a pair to prime replication of the other, thereby forming a mutagenized double-stranded polynucleotide; and

- d) repeating the second and third steps for at least two further cycles, wherein the resultant mixture in the second step of a further cycle includes the mutagenized double-stranded polynucleotide from the third step of the previous cycle, and the further cycle forms a further mutagenized double-stranded polynucleotide.

SEQUENCE LISTING

<110> Novartis

<120> Genes controlling diseases

<130> S-30431/A

<140> CGC1989

<141> 1998-03-17

<150> US 09/042763

<151> 1998-03-17

<160> 42

<170> PatentIn Ver. 2.0

<210> 1

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Conserved  
amino acid sequence

<400> 1

Glu Leu Met Xaa Xaa Gly Xaa Ile Ser Leu Leu Leu Xaa  
1 5 10

<210> 2

<211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Conserved  
amino acid sequence

<400> 2

Xaa Thr Xaa Pro Leu Xaa Xaa Xaa Val Xaa Gln Met Gly Ser  
1 5 10

<210> 3

<211> 1868

<212> DNA

<213> Triticum sp.

<220>

<221> CDS

<222> (176)..(1777)

<220>  
 <221> misc\_feature  
 <222> (365)..(403)  
 <223> location of the amino acid sequence set forth in  
 SEQ ID No: 1

<220>  
 <221> misc\_feature  
 <222> (1352)..(1393)  
 <223> location of the amino acid sequence set forth in  
 SEQ ID No: 2

<400> 3  
 cacgcccaca cttcgccaac acacaacgta cctgcgtacg tacgctttcc atttcctttc 60  
 ttgctccggc cggccggcca cgtagaatag ataccgggcc aggtagggtac ctogttgggt 120  
 cagacgaccg gcggctgggt ctccggacaa ggaaagaggt tgcgctcggg gaccg atg 178  
 Met  
 1  
 gcg gac gac gac gag tac ccc cca gcg agg acg ctg ccg gag acg ccg 226  
 Ala Asp Asp Asp Glu Tyr Pro Pro Ala Arg Thr Leu Pro Glu Thr Pro  
 5 10 15  
 tcc tgg gcg gtg gcc ctc gtc ttc gcc gtc atg atc atc gtg tcc gtc 274  
 Ser Trp Ala Val Ala Leu Val Phe Ala Val Met Ile Ile Val Ser Val  
 20 25 30  
 ctc ctg gag cac gcg ctc cat aag ctc ggc cat tgg ttc cac aag ccg 322  
 Leu Leu Glu His Ala Leu His Lys Leu Gly His Trp Phe His Lys Arg  
 35 40 45  
 cac aag aac gcg ctg gcg gag gcg ctg gag aag atc aag gcg gag ctc 370  
 His Lys Asn Ala Leu Ala Glu Ala Leu Glu Lys Ile Lys Ala Glu Leu  
 50 55 60 65  
 atg ctg gtg ggc ttc atc tcg ctg ctg ctc gcc gtg acg cag gac ccc 418  
 Met Leu Val Gly Phe Ile Ser Leu Leu Leu Ala Val Thr Gln Asp Pro  
 70 75 80  
 atc tcc ggg ata tgc atc tcc gag aag gcc gcc agc atc atg ccg ccc 466  
 Ile Ser Gly Ile Cys Ile Ser Glu Lys Ala Ala Ser Ile Met Arg Pro  
 85 90 95  
 tgc aag ctg ccc cct ggc tcc gtc aag agc aag tac aaa gac tac tac 514  
 Cys Lys Leu Pro Pro Gly Ser Val Lys Ser Lys Tyr Lys Asp Tyr Tyr  
 100 105 110  
 tgc gcc aaa cag ggc aag gtg tcg ctc atg tcc acg ggc agc ttg cac 562  
 Cys Ala Lys Gln Gly Lys Val Ser Leu Met Ser Thr Gly Ser Leu His  
 115 120 125  
 cag ctg cac ata ttc atc ttc gtg ctc gcc gtc ttc cat gtc acc tac 610  
 Gln Leu His Ile Phe Ile Phe Val Leu Ala Val Phe His Val Thr Tyr

130	135	140	145	
agc gtc atc atc atg gct cta agc cgt ctc aaa atg aga acc tgg aag				658
Ser Val Ile Ile Met Ala Leu Ser Arg Leu Lys Met Arg Thr Trp Lys	150	155	160	
aaa tgg gag aca gag acc gcc tcc ctg gaa tac cag ttc gca aat gat				706
Lys Trp Glu Thr Glu Thr Ala Ser Leu Glu Tyr Gln Phe Ala Asn Asp	165	170	175	
cct gcg cgg ttc cgc ttc acg cac cag acg tcg ttc gtg aag cgg cac				754
Pro Ala Arg Phe Arg Phe Thr His Gln Thr Ser Phe Val Lys Arg His	180	185	190	
ctg ggc ctc tcc agc acc ccc ggc gtc aga tgg gtg gtg gcc ttc ttc				802
Leu Gly Leu Ser Ser Thr Pro Gly Val Arg Trp Val Val Ala Phe Phe	195	200	205	
agg cag ttc ttc agg tcg gtc acc aag gtg gac tac ctc acc ttg agg				850
Arg Gln Phe Phe Arg Ser Val Thr Lys Val Asp Tyr Leu Thr Leu Arg	210	215	220	225
gca ggc ttc atc aac gcg cat ttg tcg cat aac agc aag ttc gac ttc				898
Ala Gly Phe Ile Asn Ala His Leu Ser His Asn Ser Lys Phe Asp Phe	230	235	240	
cac aag tac atc aag agg tcc atg gag gac gac ttc aaa gtc gtc gtt				946
His Lys Tyr Ile Lys Arg Ser Met Glu Asp Asp Phe Lys Val Val Val	245	250	255	
ggc atc agc ctc cgg ctg tgg tgt gtg gcg atc ctc acc ctc ttc ctt				994
Gly Ile Ser Leu Pro Leu Trp Cys Val Ala Ile Leu Thr Leu Phe Leu	260	265	270	
gac att gac ggg atc ggc acg ctc acc tgg att tct ttc atc cct ctc				1042
Asp Ile Asp Gly Ile Gly Thr Leu Thr Trp Ile Ser Phe Ile Pro Leu	275	280	285	
gtc atc ctc ttg tgt gtt gga acc aag ctg gag atg atc atc atg gag				1090
Val Ile Leu Leu Cys Val Gly Thr Lys Leu Glu Met Ile Ile Met Glu	290	295	300	305
atg gcc ctg gag atc cag gac cgg gcg agc gtc atc aag ggg gcg ccc				1138
Met Ala Leu Glu Ile Gln Asp Arg Ala Ser Val Ile Lys Gly Ala Pro	310	315	320	
gtg gtt gag ccc agc aac aag ttc ttc tgg ttc cac cgc ccc gac tgg				1186
Val Val Glu Pro Ser Asn Lys Phe Phe Trp Phe His Arg Pro Asp Trp	325	330	335	
gtc ctc ttc ttc ata cac ctg acg cta ttc cag aac gcg ttt cag atg				1234
Val Leu Phe Phe Ile His Leu Thr Leu Phe Gln Asn Ala Phe Gln Met	340	345	350	
gca cat ttc gtg tgg aca gtg gcc acg ccc ggc ttg aag aaa tgc ttc				1282

Ala	His	Phe	Val	Trp	Thr	Val	Ala	Thr	Pro	Gly	Leu	Lys	Lys	Cys	Phe	
355						360					365					
cat	atg	cac	atc	ggg	ctg	agc	atc	atg	aag	gtc	gtg	ctg	ggg	ctg	gct	1330
His	Met	His	Ile	Gly	Leu	Ser	Ile	Met	Lys	Val	Val	Leu	Gly	Leu	Ala	
370					375					380					385	
ctt	cag	ttc	ctc	tgc	agc	tat	atc	acc	ttc	ccg	ctc	tac	gcg	ctc	gtc	1378
Leu	Gln	Phe	Leu	Cys	Ser	Tyr	Ile	Thr	Phe	Pro	Leu	Tyr	Ala	Leu	Val	
				390					395					400		
aca	cag	atg	gga	tca	aac	atg	aag	agg	tcc	atc	ttc	gac	gag	cag	acg	1426
Thr	Gln	Met	Gly	Ser	Asn	Met	Lys	Arg	Ser	Ile	Phe	Asp	Glu	Gln	Thr	
			405					410					415			
gcc	aag	gcg	ctg	aca	aac	tgg	cgg	aac	acg	gcc	aag	gag	aag	aag	aag	1474
Ala	Lys	Ala	Leu	Thr	Asn	Trp	Arg	Asn	Thr	Ala	Lys	Glu	Lys	Lys	Lys	
			420					425					430			
gtc	cga	gac	acg	gac	atg	ctg	atg	gcg	cag	atg	atc	ggc	gac	gcg	acg	1522
Val	Arg	Asp	Thr	Asp	Met	Leu	Met	Ala	Gln	Met	Ile	Gly	Asp	Ala	Thr	
			435				440					445				
ccc	agc	cga	ggg	gcg	tgc	ccc	atg	cct	agc	cgg	ggc	tcg	tcg	cca	gtg	1570
Pro	Ser	Arg	Gly	Ala	Ser	Pro	Met	Pro	Ser	Arg	Gly	Ser	Ser	Pro	Val	
					455					460					465	
cac	ctg	ctt	cac	aag	ggc	atg	gga	cgg	tcc	gac	gat	ccc	cag	agc	acg	1618
His	Leu	Leu	His	Lys	Gly	Met	Gly	Arg	Ser	Asp	Asp	Pro	Gln	Ser	Thr	
				470					475					480		
cca	acc	tcg	cca	agg	gcc	atg	gag	gag	gct	agg	gac	atg	tac	ccg	gtt	1666
Pro	Thr	Ser	Pro	Arg	Ala	Met	Glu	Glu	Ala	Arg	Asp	Met	Tyr	Pro	Val	
				485				490					495			
gtg	gtg	gcg	cat	cca	gtg	cac	aga	cta	aat	cct	gct	gac	agg	aga	agg	1714
Val	Val	Ala	His	Pro	Val	His	Arg	Leu	Asn	Pro	Ala	Asp	Arg	Arg	Arg	
				500				505				510				
tcg	gtc	tcg	tcg	tcg	gca	ctc	gat	gtc	gac	att	ccc	agc	gca	gat	ttt	1762
Ser	Val	Ser	Ser	Ser	Ala	Leu	Asp	Val	Asp	Ile	Pro	Ser	Ala	Asp	Phe	
				515			520				525					
tcc	ttc	agc	cag	gga	tgagacaagt	ttctgtattg	atgttagtcc	aatgtatagc								1817
Ser	Phe	Ser	Gln	Gly												
530																
caacatagga	tgatcatgatt	cgtacaataa	gaaatacaaa	tttttactga	g											1868

&lt;210&gt; 4

&lt;211&gt; 534

&lt;212&gt; PRT

&lt;213&gt; Triticum sp.



&lt;400&gt; 4

```

Met Ala Asp Asp Asp Glu Tyr Pro Pro Ala Arg Thr Leu Pro Glu Thr
 1           5           10           15

Pro Ser Trp Ala Val Ala Leu Val Phe Ala Val Met Ile Ile Val Ser
      20           25           30

Val Leu Leu Glu His Ala Leu His Lys Leu Gly His Trp Phe His Lys
      35           40           45

Arg His Lys Asn Ala Leu Ala Glu Ala Leu Glu Lys Ile Lys Ala Glu
      50           55           60

Leu Met Leu Val Gly Phe Ile Ser Leu Leu Leu Ala Val Thr Gln Asp
      65           70           75           80

Pro Ile Ser Gly Ile Cys Ile Ser Glu Lys Ala Ala Ser Ile Met Arg
      85           90           95

Pro Cys Lys Leu Pro Pro Gly Ser Val Lys Ser Lys Tyr Lys Asp Tyr
      100          105          110

Tyr Cys Ala Lys Gln Gly Lys Val Ser Leu Met Ser Thr Gly Ser Leu
      115          120          125

His Gln Leu His Ile Phe Ile Phe Val Leu Ala Val Phe His Val Thr
      130          135          140

Tyr Ser Val Ile Ile Met Ala Leu Ser Arg Leu Lys Met Arg Thr Trp
      145          150          155          160

Lys Lys Trp Glu Thr Glu Thr Ala Ser Leu Glu Tyr Gln Phe Ala Asn
      165          170          175

Asp Pro Ala Arg Phe Arg Phe Thr His Gln Thr Ser Phe Val Lys Arg
      180          185          190

His Leu Gly Leu Ser Ser Thr Pro Gly Val Arg Trp Val Val Ala Phe
      195          200          205

Phe Arg Gln Phe Phe Arg Ser Val Thr Lys Val Asp Tyr Leu Thr Leu
      210          215          220

Arg Ala Gly Phe Ile Asn Ala His Leu Ser His Asn Ser Lys Phe Asp
      225          230          235          240

Phe His Lys Tyr Ile Lys Arg Ser Met Glu Asp Asp Phe Lys Val Val
      245          250          255

Val Gly Ile Ser Leu Pro Leu Trp Cys Val Ala Ile Leu Thr Leu Phe
      260          265          270

Leu Asp Ile Asp Gly Ile Gly Thr Leu Thr Trp Ile Ser Phe Ile Pro
      275          280          285

```

Leu Val Ile Leu Leu Cys Val Gly Thr Lys Leu Glu Met Ile Ile Met  
 290 295 300  
 Glu Met Ala Leu Glu Ile Gln Asp Arg Ala Ser Val Ile Lys Gly Ala  
 305 310 315 320  
 Pro Val Val Glu Pro Ser Asn Lys Phe Phe Trp Phe His Arg Pro Asp  
 325 330 335  
 Trp Val Leu Phe Phe Ile His Leu Thr Leu Phe Gln Asn Ala Phe Gln  
 340 345 350  
 Met Ala His Phe Val Trp Thr Val Ala Thr Pro Gly Leu Lys Lys Cys  
 355 360 365  
 Phe His Met His Ile Gly Leu Ser Ile Met Lys Val Val Leu Gly Leu  
 370 375 380  
 Ala Leu Gln Phe Leu Cys Ser Tyr Ile Thr Phe Pro Leu Tyr Ala Leu  
 385 390 395 400  
 Val Thr Gln Met Gly Ser Asn Met Lys Arg Ser Ile Phe Asp Glu Gln  
 405 410 415  
 Thr Ala Lys Ala Leu Thr Asn Trp Arg Asn Thr Ala Lys Glu Lys Lys  
 420 425 430  
 Lys Val Arg Asp Thr Asp Met Leu Met Ala Gln Met Ile Gly Asp Ala  
 435 440 445  
 Thr Pro Ser Arg Gly Ala Ser Pro Met Pro Ser Arg Gly Ser Ser Pro  
 450 455 460  
 Val His Leu Leu His Lys Gly Met Gly Arg Ser Asp Asp Pro Gln Ser  
 465 470 475 480  
 Thr Pro Thr Ser Pro Arg Ala Met Glu Glu Ala Arg Asp Met Tyr Pro  
 485 490 495  
 Val Val Val Ala His Pro Val His Arg Leu Asn Pro Ala Asp Arg Arg  
 500 505 510  
 Arg Ser Val Ser Ser Ser Ala Leu Asp Val Asp Ile Pro Ser Ala Asp  
 515 520 525  
 Phe Ser Phe Ser Gln Gly  
 530

&lt;210&gt; 5

&lt;211&gt; 1693

&lt;212&gt; DNA

&lt;213&gt; Triticum sp.

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1602)

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (190)..(228)

<223> location of the amino acid sequence set forth in  
SEQ ID No:1

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1177)..(1218)

<223> location of the amino acid sequence set forth in  
SEQ ID No:2

&lt;400&gt; 5

atg gcg gag gac tac gag tac ccc ccg gcg cgg acg ctg ccg gag acg 48

Met Ala Glu Asp Tyr Glu Tyr Pro Pro Ala Arg Thr Leu Pro Glu Thr  
1 5 10 15

ccg tcc tgg gcg gtg gcg ctc gtc ttc gcc gtc atg atc atc gtg tcc 96

Pro Ser Trp Ala Val Ala Leu Val Phe Ala Val Met Ile Ile Val Ser  
20 25 30

gtc ctc ctg gag cac gcg ctc cac aag ctc ggc cat tgg ttc cac aag 144

Val Leu Leu Glu His Ala Leu His Lys Leu Gly His Trp Phe His Lys  
35 40 45

cgg cac aag aac gcg ctg gcg gag gcg ctg gag aag atc aaa gcg gag 192

Arg His Lys Asn Ala Leu Ala Glu Ala Leu Glu Lys Ile Lys Ala Glu  
50 55 60

ctg atg ctg gtg ggg ttc atc tcg ctg ctg ctc gcc gtg acg cag gac 240

Leu Met Leu Val Gly Phe Ile Ser Leu Leu Ala Val Thr Gln Asp  
65 70 75 80

cca atc tcc ggg ata tgc atc tcc gag aag gcc gcc agc atc atg cgg 288

Pro Ile Ser Gly Ile Cys Ile Ser Glu Lys Ala Ala Ser Ile Met Arg  
85 90 95

ccc tgc agc ctg ccc cct ggt tcc gtc aag agc aag tac aaa gac tac 336

Pro Cys Ser Leu Pro Pro Gly Ser Val Lys Ser Lys Tyr Lys Asp Tyr  
100 105 110

tac tgc gcc aaa aag ggc aag gtg tcg cta atg tcc acg ggc agc ttg 384

Tyr Cys Ala Lys Lys Gly Lys Val Ser Leu Met Ser Thr Gly Ser Leu  
115 120 125

cac cag ctc cac atg ttc atc ttc gtg ctc gcc gtc ttc cat gtc acc 432

His Gln Leu His Met Phe Ile Phe Val Leu Ala Val Phe His Val Thr  
130 135 140

tac agc gtc atc atc atg gct cta agc cgt ctc aaa atg agg aca tgg 480

Tyr Ser Val Ile Ile Met Ala Leu Ser Arg Leu Lys Met Arg Thr Trp  
145 150 155 160

aag aaa tgg gag aca gag acc gyc tcc ttg gaa tac cag ttc gca aat	528
Lys Lys Trp Glu Thr Glu Thr Xaa Ser Leu Glu Tyr Gln Phe Ala Asn	
165 170 175	
gat cct gcg cgg ttc cgc ttc acg cac cag acg tcg ttc gtg aag cgt	576
Asp Pro Ala Arg Phe Arg Phe Thr His Gln Thr Ser Phe Val Lys Arg	
180 185 190	
cac ctg ggc ctc tcc agc acc ccc ggc atc aga tgg gtg gtg gcc ttc	624
His Leu Gly Leu Ser Ser Thr Pro Gly Ile Arg Trp Val Val Ala Phe	
195 200 205	
ttc agg cag ttc ttc agg tcg gtc acc aag gtg gac tac ctc acc ctg	672
Phe Arg Gln Phe Phe Arg Ser Val Thr Lys Val Asp Tyr Leu Thr Leu	
210 215 220	
agg gca ggc ttc atc aac gcg cat ttg tcg cat aac agc aag ttc gac	720
Arg Ala Gly Phe Ile Asn Ala His Leu Ser His Asn Ser Lys Phe Asp	
225 230 235 240	
ttc cac aag tac atc aag agg tcc atg gag gac gac ttc aaa gtc gtc	768
Phe His Lys Tyr Ile Lys Arg Ser Met Glu Asp Asp Phe Lys Val Val	
245 250 255	
gtt ggc atc agc ctc ccg ctg tgg tgt gtg gcg atc ctc acc ctc ttc	816
Val Gly Ile Ser Leu Pro Leu Trp Cys Val Ala Ile Leu Thr Leu Phe	
260 265 270	
ctt gat att gac ggg atc ggc acg ctc acc tgg att tct ttc atc cct	864
Leu Asp Ile Asp Gly Ile Gly Thr Leu Thr Trp Ile Ser Phe Ile Pro	
275 280 285	
ctc gtc atc ctc ttg tgt gtt gga acc aag ctg gag atg atc atc atg	912
Leu Val Ile Leu Leu Cys Val Gly Thr Lys Leu Glu Met Ile Ile Met	
290 295 300	
gag atg gcc ctg gag atc cag gac cgg gcg agc gtc atc aag ggg gcg	960
Glu Met Ala Leu Glu Ile Gln Asp Arg Ala Ser Val Ile Lys Gly Ala	
305 310 315 320	
ccc gtg gtt gag ccc agc aac aag ttc ttc tgg ttc cac cgc ccc gac	1008
Pro Val Val Glu Pro Ser Asn Lys Phe Phe Trp Phe His Arg Pro Asp	
325 330 335	
tgg gtc ctc ttc ttc ata cac ctg acg ctg ttc cag aat gcg ttt cag	1056
Trp Val Leu Phe Phe Ile His Leu Thr Leu Phe Gln Asn Ala Phe Gln	
340 345 350	
atg gca cat ttc gtc tgg aca gtg gcc acg ccc ggc ttg aag aaa tgc	1104
Met Ala His Phe Val Trp Thr Val Ala Thr Pro Gly Leu Lys Lys Cys	
355 360 365	
ttc cat atg cac atc ggt ctg agc atc atg aag gtc gtg ctg ggg ctg	1152
Phe His Met His Ile Gly Leu Ser Ile Met Lys Val Val Leu Gly Leu	

370	375	380	
gct ctt cag ttc ctc tgc agc tat atc acc ttc ccc ctc tac gcg ctc			1200
Ala Leu Gln Phe Leu Cys Ser Tyr Ile Thr Phe Pro Leu Tyr Ala Leu			
385	390	395	400
gtc aca cag atg gga tgc aac atg aag agg tcc atc ttc gac gag cag			1248
Val Thr Gln Met Gly Ser Asn Met Lys Arg Ser Ile Phe Asp Glu Gln			
	405	410	415
acg gcc aag gcg ctg acc aac tgg cgg aac acg gcc aag gag aag aag			1296
Thr Ala Lys Ala Leu Thr Asn Trp Arg Asn Thr Ala Lys Glu Lys Lys			
	420	425	430
aag gtc cga gac acg gac atg ctg atg gcg cag atg atc ggc gac gcg			1344
Lys Val Arg Asp Thr Asp Met Leu Met Ala Gln Met Ile Gly Asp Ala			
	435	440	445
acg ccc agc cga ggc acg tgc cgg atg cct agc cgg gct tgc tca ccg			1392
Thr Pro Ser Arg Gly Thr Ser Pro Met Pro Ser Arg Ala Ser Ser Pro			
	450	455	460
gtg cac ctg ctt cac aag ggc atg gga cgg tcc gac gat ccc cag agc			1440
Val His Leu Leu His Lys Gly Met Gly Arg Ser Asp Asp Pro Gln Ser			
	465	470	475
gcg ccg acc tgc cca agg acc atg gag gag gct agg gac atg tac ccg			1488
Ala Pro Thr Ser Pro Arg Thr Met Glu Glu Ala Arg Asp Met Tyr Pro			
	485	490	495
gtt gtg gtg gcg cat ccc gtg cac aga cta aat cct gct gac agg cgg			1536
Val Val Val Ala His Pro Val His Arg Leu Asn Pro Ala Asp Arg Arg			
	500	505	510
agg tgc gtc tct tgc tgc gca ctc gat gcc gac atc ccc agc gca gat			1584
Arg Ser Val Ser Ser Ser Ala Leu Asp Ala Asp Ile Pro Ser Ala Asp			
	515	520	525
ttt tcc ttc agc cag gga tgagacaagt ttctgtattg atgttagtcc			1632
Phe Ser Phe Ser Gln Gly			
	530		
aatgtatagc caacatagga tgtcatgatt cgtacaataa gaaatacaaa tttttactga			1692
g			1693

&lt;210&gt; 6

&lt;211&gt; 534

&lt;212&gt; PRT

&lt;213&gt; Triticum sp.

&lt;400&gt; 6

Met	Ala	Glu	Asp	Tyr	Glu	Tyr	Pro	Pro	Ala	Arg	Thr	Leu	Pro	Glu	Thr
1					5				10					15	

Pro Ser Trp Ala Val Ala Leu Val Phe Ala Val Met Ile Ile Val Ser  
                   20                                  25                                  30  
 Val Leu Leu Glu His Ala Leu His Lys Leu Gly His Trp Phe His Lys  
                   35                                  40                                  45  
 Arg His Lys Asn Ala Leu Ala Glu Ala Leu Glu Lys Ile Lys Ala Glu  
                   50                                  55                                  60  
 Leu Met Leu Val Gly Phe Ile Ser Leu Leu Leu Ala Val Thr Gln Asp  
                   65                                  70                                  75                                  80  
 Pro Ile Ser Gly Ile Cys Ile Ser Glu Lys Ala Ala Ser Ile Met Arg  
                                   85                                  90                                  95  
 Pro Cys Ser Leu Pro Pro Gly Ser Val Lys Ser Lys Tyr Lys Asp Tyr  
                                   100                                  105                                  110  
 Tyr Cys Ala Lys Lys Gly Lys Val Ser Leu Met Ser Thr Gly Ser Leu  
                   115                                  120                                  125  
 His Gln Leu His Met Phe Ile Phe Val Leu Ala Val Phe His Val Thr  
                   130                                  135                                  140  
 Tyr Ser Val Ile Ile Met Ala Leu Ser Arg Leu Lys Met Arg Thr Trp  
                   145                                  150                                  155                                  160  
 Lys Lys Trp Glu Thr Glu Thr Xaa Ser Leu Glu Tyr Gln Phe Ala Asn  
                                   165                                  170                                  175  
 Asp Pro Ala Arg Phe Arg Phe Thr His Gln Thr Ser Phe Val Lys Arg  
                                   180                                  185                                  190  
 His Leu Gly Leu Ser Ser Thr Pro Gly Ile Arg Trp Val Val Ala Phe  
                   195                                  200                                  205  
 Phe Arg Gln Phe Phe Arg Ser Val Thr Lys Val Asp Tyr Leu Thr Leu  
                   210                                  215                                  220  
 Arg Ala Gly Phe Ile Asn Ala His Leu Ser His Asn Ser Lys Phe Asp  
                   225                                  230                                  235                                  240  
 Phe His Lys Tyr Ile Lys Arg Ser Met Glu Asp Asp Phe Lys Val Val  
                                   245                                  250                                  255  
 Val Gly Ile Ser Leu Pro Leu Trp Cys Val Ala Ile Leu Thr Leu Phe  
                   260                                  265                                  270  
 Leu Asp Ile Asp Gly Ile Gly Thr Leu Thr Trp Ile Ser Phe Ile Pro  
                   275                                  280                                  285  
 Leu Val Ile Leu Leu Cys Val Gly Thr Lys Leu Glu Met Ile Ile Met  
                   290                                  295                                  300

Glu Met Ala Leu Glu Ile Gln Asp Arg Ala Ser Val Ile Lys Gly Ala  
 305 310 315 320  
 Pro Val Val Glu Pro Ser Asn Lys Phe Phe Trp Phe His Arg Pro Asp  
 325 330 335  
 Trp Val Leu Phe Phe Ile His Leu Thr Leu Phe Gln Asn Ala Phe Gln  
 340 345 350  
 Met Ala His Phe Val Trp Thr Val Ala Thr Pro Gly Leu Lys Lys Cys  
 355 360 365  
 Phe His Met His Ile Gly Leu Ser Ile Met Lys Val Val Leu Gly Leu  
 370 375 380  
 Ala Leu Gln Phe Leu Cys Ser Tyr Ile Thr Phe Pro Leu Tyr Ala Leu  
 385 390 395 400  
 Val Thr Gln Met Gly Ser Asn Met Lys Arg Ser Ile Phe Asp Glu Gln  
 405 410 415  
 Thr Ala Lys Ala Leu Thr Asn Trp Arg Asn Thr Ala Lys Glu Lys Lys  
 420 425 430  
 Lys Val Arg Asp Thr Asp Met Leu Met Ala Gln Met Ile Gly Asp Ala  
 435 440 445  
 Thr Pro Ser Arg Gly Thr Ser Pro Met Pro Ser Arg Ala Ser Ser Pro  
 450 455 460  
 Val His Leu Leu His Lys Gly Met Gly Arg Ser Asp Asp Pro Gln Ser  
 465 470 475 480  
 Ala Pro Thr Ser Pro Arg Thr Met Glu Glu Ala Arg Asp Met Tyr Pro  
 485 490 495  
 Val Val Val Ala His Pro Val His Arg Leu Asn Pro Ala Asp Arg Arg  
 500 505 510  
 Arg Ser Val Ser Ser Ser Ala Leu Asp Ala Asp Ile Pro Ser Ala Asp  
 515 520 525  
 Phe Ser Phe Ser Gln Gly  
 530

&lt;210&gt; 7

&lt;211&gt; 1886

&lt;212&gt; DNA

&lt;213&gt; Triticum sp.

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (198) .. (1799)

[illegible]



gtc ttc cat gtc acc tac agc gtc atc atc atg gct cta agc cgt ctc	662
Val Phe His Val Thr Tyr Ser Val Ile Ile Met Ala Leu Ser Arg Leu	
140 145 150 155	
aag atg aga aca tgg aag aaa tgg gag aca gaa acc gcc tcc ttg gaa	710
Lys Met Arg Thr Trp Lys Lys Trp Glu Thr Glu Thr Ala Ser Leu Glu	
160 165 170	
tac cag ttc gca aat gat cct gcg cgg ttc cgc ttc acg cac cag acg	758
Tyr Gln Phe Ala Asn Asp Pro Ala Arg Phe Arg Phe Thr His Gln Thr	
175 180 185	
tcg ttc gtg aag cgg cac ctg gcc ctg tcc agc acc ccc gcc gtc aga	806
Ser Phe Val Lys Arg His Leu Gly Leu Ser Ser Thr Pro Gly Val Arg	
190 195 200	
tgg gtg gtg gcc ttc ttc agg cag ttc ttc agg tcg gtc acc aag gtg	854
Trp Val Val Ala Phe Phe Arg Gln Phe Phe Arg Ser Val Thr Lys Val	
205 210 215	
gac tac ctc acc ttg agg gca gcc ttc atc aac gcg cac ttg tcg cag	902
Asp Tyr Leu Thr Leu Arg Ala Gly Phe Ile Asn Ala His Leu Ser Gln	
220 225 230 235	
aac agc aag ttc gac ttc cac aag tac atc aag agg tcc atg gag gac	950
Asn Ser Lys Phe Asp Phe His Lys Tyr Ile Lys Arg Ser Met Glu Asp	
240 245 250	
gac ttc aaa gtc gtc gtt gcc atc agc ctc ccg ctg tgg gct gtg gcg	998
Asp Phe Lys Val Val Val Gly Ile Ser Leu Pro Leu Trp Ala Val Ala	
255 260 265	
atc ctc acc ctc ttc ctt gat atc gac ggg atc ggc aca ctc acc tgg	1046
Ile Leu Thr Leu Phe Leu Asp Ile Asp Gly Ile Gly Thr Leu Thr Trp	
270 275 280	
gtt tct ttc atc cct ctc atc atc ctc ttg tgt gtt gga acc aag cta	1094
Val Ser Phe Ile Pro Leu Ile Ile Leu Leu Cys Val Gly Thr Lys Leu	
285 290 295	
gag atg atc atc atg ggg atg gcc ctg gag atc cag gac cgg tcg agc	1142
Glu Met Ile Ile Met Gly Met Ala Leu Glu Ile Gln Asp Arg Ser Ser	
300 305 310 315	
gtc atc aag ggg gca ccc gtg gtc gag ccc agc aac aag ttc ttc tgg	1190
Val Ile Lys Gly Ala Pro Val Val Glu Pro Ser Asn Lys Phe Phe Trp	
320 325 330	
ttc cac cgc ccc gac tgg gtc ctc ttc ttc ata cac ctg acg ctg ttc	1238
Phe His Arg Pro Asp Trp Val Leu Phe Phe Ile His Leu Thr Leu Phe	
335 340 345	
cag aac gcg ttt cag atg gca cat ttc gtg tgg aca gtg gcc acg ccc	1286
Gln Asn Ala Phe Gln Met Ala His Phe Val Trp Thr Val Ala Thr Pro	
350 355 360	

ggc ttg aag gac tgc ttc cat atg aac atc ggg ctg agc atc atg aag 1334  
 Gly Leu Lys Asp Cys Phe His Met Asn Ile Gly Leu Ser Ile Met Lys  
 365 370 375  
 gtc gtg ctg ggg ctg gct ctc cag ttc ctg tgc agc tac atc acc ttc 1382  
 Val Val Leu Gly Leu Ala Leu Gln Phe Leu Cys Ser Tyr Ile Thr Phe  
 380 385 390 395  
 ccc ctc tac gcg cta gtc aca cag atg gga tca aac atg aag agg tcc 1430  
 Pro Leu Tyr Ala Leu Val Thr Gln Met Gly Ser Asn Met Lys Arg Ser  
 400 405 410  
 atc ttc gac gag cag aca gcc aag gcg ctg acc aac tgg cgg aac acg 1478  
 Ile Phe Asp Glu Gln Thr Ala Lys Ala Leu Thr Asn Trp Arg Asn Thr  
 415 420 425  
 gcc aag gag aag aag aag gtc cga gac acg gac atg ctg atg gcg cag 1526  
 Ala Lys Glu Lys Lys Lys Val Arg Asp Thr Asp Met Leu Met Ala Gln  
 430 435 440  
 atg atc ggc gac gca aca ccc agc cga ggc acg tcc ccg atg cct agc 1574  
 Met Ile Gly Asp Ala Thr Pro Ser Arg Gly Thr Ser Pro Met Pro Ser  
 445 450 455  
 cgg ggc tca tcg ccg gtg cac ctg ctt cag aag ggc atg gga cgg tct 1622  
 Arg Gly Ser Ser Pro Val His Leu Leu Gln Lys Gly Met Gly Arg Ser  
 460 465 470 475  
 gac gat ccc cag agc gca ccg acc tcg cca agg acc atg gag gag gct 1670  
 Asp Asp Pro Gln Ser Ala Pro Thr Ser Pro Arg Thr Met Glu Glu Ala  
 480 485 490  
 agg gac atg tac ccg gtt gtg gtg gcg cat cct gta cac aga cta aat 1718  
 Arg Asp Met Tyr Pro Val Val Val Ala His Pro Val His Arg Leu Asn  
 495 500 505  
 cct gct gac agg cgg agg tcg gtc tct tca tca gcc ctc gat gcc gac 1766  
 Pro Ala Asp Arg Arg Arg Ser Val Ser Ser Ser Ala Leu Asp Ala Asp  
 510 515 520  
 atc ccc agc gca gat ttt tcc ttc agc cag gga tgagacaagt ttctgtattg 1819  
 Ile Pro Ser Ala Asp Phe Ser Phe Ser Gln Gly  
 525 530  
 atgttagtcc aatgtatagc caacatagga tgtgatgatt cgtacaataa gaaatacaat 1879  
 tttttac 1886

&lt;210&gt; 8

&lt;211&gt; 534

&lt;212&gt; PRT

&lt;213&gt; Triticum sp.

&lt;400&gt; 8

```

Met Ala Lys Asp Asp Gly Tyr Pro Pro Ala Arg Thr Leu Pro Glu Thr
 1           5           10           15

Pro Ser Trp Ala Val Ala Leu Val Phe Ala Val Met Ile Ile Val Ser
      20           25           30

Val Leu Leu Glu His Ala Leu His Lys Leu Gly His Trp Phe His Lys
 35           40           45

Arg His Lys Asn Ala Leu Ala Glu Ala Leu Glu Lys Met Lys Ala Glu
 50           55           60

Leu Met Leu Val Gly Phe Ile Ser Leu Leu Leu Ala Val Thr Gln Asp
 65           70           75           80

Pro Ile Ser Gly Ile Cys Ile Ser Gln Lys Ala Ala Ser Ile Met Arg
      85           90           95

Pro Cys Lys Val Glu Pro Gly Ser Val Lys Ser Lys Tyr Lys Asp Tyr
      100          105          110

Tyr Cys Ala Lys Glu Gly Lys Val Ala Leu Met Ser Thr Gly Ser Leu
      115          120          125

His Gln Leu His Ile Phe Ile Phe Val Leu Ala Val Phe His Val Thr
      130          135          140

Tyr Ser Val Ile Ile Met Ala Leu Ser Arg Leu Lys Met Arg Thr Trp
      145          150          155          160

Lys Lys Trp Glu Thr Glu Thr Ala Ser Leu Glu Tyr Gln Phe Ala Asn
      165          170          175

Asp Pro Ala Arg Phe Arg Phe Thr His Gln Thr Ser Phe Val Lys Arg
      180          185          190

His Leu Gly Leu Ser Ser Thr Pro Gly Val Arg Trp Val Val Ala Phe
      195          200          205

Phe Arg Gln Phe Phe Arg Ser Val Thr Lys Val Asp Tyr Leu Thr Leu
      210          215          220

Arg Ala Gly Phe Ile Asn Ala His Leu Ser Gln Asn Ser Lys Phe Asp
      225          230          235          240

Phe His Lys Tyr Ile Lys Arg Ser Met Glu Asp Asp Phe Lys Val Val
      245          250          255

Val Gly Ile Ser Leu Pro Leu Trp Ala Val Ala Ile Leu Thr Leu Phe
      260          265          270

Leu Asp Ile Asp Gly Ile Gly Thr Leu Thr Trp Val Ser Phe Ile Pro
      275          280          285

```

Leu Ile Ile Leu Leu Cys Val Gly Thr Lys Leu Glu Met Ile Ile Met  
 290 295 300  
 Gly Met Ala Leu Glu Ile Gln Asp Arg Ser Ser Val Ile Lys Gly Ala  
 305 310 315 320  
 Pro Val Val Glu Pro Ser Asn Lys Phe Phe Trp Phe His Arg Pro Asp  
 325 330 335  
 Trp Val Leu Phe Phe Ile His Leu Thr Leu Phe Gln Asn Ala Phe Gln  
 340 345 350  
 Met Ala His Phe Val Trp Thr Val Ala Thr Pro Gly Leu Lys Asp Cys  
 355 360 365  
 Phe His Met Asn Ile Gly Leu Ser Ile Met Lys Val Val Leu Gly Leu  
 370 375 380  
 Ala Leu Gln Phe Leu Cys Ser Tyr Ile Thr Phe Pro Leu Tyr Ala Leu  
 385 390 395 400  
 Val Thr Gln Met Gly Ser Asn Met Lys Arg Ser Ile Phe Asp Glu Gln  
 405 410 415  
 Thr Ala Lys Ala Leu Thr Asn Trp Arg Asn Thr Ala Lys Glu Lys Lys  
 420 425 430  
 Lys Val Arg Asp Thr Asp Met Leu Met Ala Gln Met Ile Gly Asp Ala  
 435 440 445  
 Thr Pro Ser Arg Gly Thr Ser Pro Met Pro Ser Arg Gly Ser Ser Pro  
 450 455 460  
 Val His Leu Leu Gln Lys Gly Met Gly Arg Ser Asp Asp Pro Gln Ser  
 465 470 475 480  
 Ala Pro Thr Ser Pro Arg Thr Met Glu Glu Ala Arg Asp Met Tyr Pro  
 485 490 495  
 Val Val Val Ala His Pro Val His Arg Leu Asn Pro Ala Asp Arg Arg  
 500 505 510  
 Arg Ser Val Ser Ser Ser Ala Leu Asp Ala Asp Ile Pro Ser Ala Asp  
 515 520 525  
 Phe Ser Phe Ser Gln Gly  
 530

&lt;210&gt; 9

&lt;211&gt; 2197

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

<221> CDS  
<222> (331)..(2037)

<220>  
<221> misc\_feature  
<222> (589)..(627)  
<223> location of the amino acid sequence set forth in  
SEQ ID No:1

<220>  
<221> misc\_feature  
<222> (1603)..(1644)  
<223> location of the amino acid sequence set forth in  
SEQ ID No:2

<400> 9  
agtaatttag ctgttcttct acccctctga tctctcacag gggatcaaat agttttgata 60  
catagagcca caacagtgc attagtgtgt tgttgactac tgtaagggtt gggttttgaa 120  
aagagacatg aaggagtgtt attaggttga ttgtcttcaa gtacctccag tgtcaaacaa 180  
acattgacga ttgattctct tcccataatt tattgttatg cattacatat cacagtaaac 240  
ggactttcaa gtcaacaccg catttatttg cctcttcat tgtttcacgt acgtaatcaa 300  
ggaccaaggg attttgttct tttggctacc atg gcc aca aga tgc ttt tgg tgt 354  
Met Ala Thr Arg Cys Phe Trp Cys  
1 5  
tgg acc act ttg ctc ttc tgc tct cag ctg ctt acc ggc ttt gcc cga 402  
Trp Thr Thr Leu Leu Phe Cys Ser Gln Leu Leu Thr Gly Phe Ala Arg  
10 15 20  
gct tcc tct gca ggc ggc gcc aaa gag aaa gga ctc tcc caa act ccc 450  
Ala Ser Ser Ala Gly Gly Ala Lys Glu Lys Gly Leu Ser Gln Thr Pro  
25 30 35 40  
acc tgg gcc gtt gcc ctc gtc tgt acc ttt ttc att ctt gtc tcc gtc 498  
Thr Trp Ala Val Ala Leu Val Cys Thr Phe Phe Ile Leu Val Ser Val  
45 50 55  
ctt ctc gag aag gct ctt cac aga gtt gcc acg tgg ttg tgg gag aaa 546  
Leu Leu Glu Lys Ala Leu His Arg Val Ala Thr Trp Leu Trp Glu Lys  
60 65 70  
cat aag aac tct ctg ctt gaa gcc ttg gaa aaa ata aag gcc gag ctg 594  
His Lys Asn Ser Leu Leu Glu Ala Leu Glu Lys Ile Lys Ala Glu Leu  
75 80 85  
atg att cta gga ttc att tcc ttg ttg cta acc ttc gga gag cag tac 642  
Met Ile Leu Gly Phe Ile Ser Leu Leu Leu Thr Phe Gly Glu Gln Tyr  
90 95 100  
att ctc aag att tgt att cct gaa aag gct gca gcc tct atg tta cct 690

Ile Leu Lys Ile Cys Ile Pro Glu Lys Ala Ala Ala Ser Met Leu Pro	
105 110 115 120	
tgt cca gct cct tct act cat gac caa gac aag acc cac cgc aga cgt	738
Cys Pro Ala Pro Ser Thr His Asp Gln Asp Lys Thr His Arg Arg Arg	
125 130 135	
cta gct gct gct acg acc tct tcc cgc tgc gat gag ggt cat gaa cca	786
Leu Ala Ala Ala Thr Thr Ser Ser Arg Cys Asp Glu Gly His Glu Pro	
140 145 150	
ctc ata cct gcc acg ggt ttg cac cag cta cac att cta ttg ttc ttc	834
Leu Ile Pro Ala Thr Gly Leu His Gln Leu His Ile Leu Leu Phe Phe	
155 160 165	
atg gct gcc ttt cat atc ctc tac agt ttc atc acc atg atg ctt ggc	882
Met Ala Ala Phe His Ile Leu Tyr Ser Phe Ile Thr Met Met Leu Gly	
170 175 180	
aga ctc aag atc cgt ggc tgg aaa aag tgg gag cag gag aca tgt tct	930
Arg Leu Lys Ile Arg Gly Trp Lys Lys Trp Glu Gln Glu Thr Cys Ser	
185 190 195 200	
cat gat tac gag ttt tca atc gat cca tca aga ttc aga ctc act cat	978
His Asp Tyr Glu Phe Ser Ile Asp Pro Ser Arg Phe Arg Leu Thr His	
205 210 215	
gag acg tcc ttt gtt aga caa cat tcc agt ttc tgg aca aaa atc ccc	1026
Glu Thr Ser Phe Val Arg Gln His Ser Ser Phe Trp Thr Lys Ile Pro	
220 225 230	
ttc ttc ttt tat gct ggg tgc ttc cta cag cag ttt ttc cga tct gtc	1074
Phe Phe Phe Tyr Ala Gly Cys Phe Leu Gln Gln Phe Phe Arg Ser Val	
235 240 245	
ggg agg act gac tac tta act ctg cgc cat ggc ttc atc gct gcc cat	1122
Gly Arg Thr Asp Tyr Leu Thr Leu Arg His Gly Phe Ile Ala Ala His	
250 255 260	
tta gct cca gga aga aag ttc gac ttc cag aag tat atc aaa aga tca	1170
Leu Ala Pro Gly Arg Lys Phe Asp Phe Gln Lys Tyr Ile Lys Arg Ser	
265 270 275 280	
ttg gaa gac gat ttc aag gtg gta gtt gga ata agt cct ctt ttg tgg	1218
Leu Glu Asp Asp Phe Lys Val Val Val Gly Ile Ser Pro Leu Leu Trp	
285 290 295	
gca tca ttt gta att ttc cta ctt ctg aat gtt aat ggc tgg gaa gca	1266
Ala Ser Phe Val Ile Phe Leu Leu Leu Asn Val Asn Gly Trp Glu Ala	
300 305 310	
ttg ttt tgg gcg tca atc cta cct gta ctt atc att cta gct gtc agt	1314
Leu Phe Trp Ala Ser Ile Leu Pro Val Leu Ile Ile Leu Ala Val Ser	
315 320 325	

acg aag ctt caa gcg atc cta aca aga atg gct ctg gga atc acg gag	1362
Thr Lys Leu Gln Ala Ile Leu Thr Arg Met Ala Leu Gly Ile Thr Glu	
330 335 340	
aga cac gca gtt gtt caa ggg ata cct ctc gtg cat ggt tca gat aag	1410
Arg His Ala Val Val Gln Gly Ile Pro Leu Val His Gly Ser Asp Lys	
345 350 355 360	
tac ttt tgg ttt aat cgc cct cag ttg cta ctt cat ctt ctt cac ttc	1458
Tyr Phe Trp Phe Asn Arg Pro Gln Leu Leu Leu His Leu Leu His Phe	
365 370 375	
gcc tta ttt cag aat gct ttc cag cta aca tac ttc ttc tgg gtc tgg	1506
Ala Leu Phe Gln Asn Ala Phe Gln Leu Thr Tyr Phe Phe Trp Val Trp	
380 385 390	
tat tcc ttt ggg cta aaa tct tgc ttt cac acg gat ttc aaa cta gtc	1554
Tyr Ser Phe Gly Leu Lys Ser Cys Phe His Thr Asp Phe Lys Leu Val	
395 400 405	
atc gta aaa ctc tct cta ggc gtt gga gct ttg att ttg tgc agc tac	1602
Ile Val Lys Leu Ser Leu Gly Val Gly Ala Leu Ile Leu Cys Ser Tyr	
410 415 420	
atc aca ctt cct ttg tat gca cta gtt act cag atg ggt tca aac atg	1650
Ile Thr Leu Pro Leu Tyr Ala Leu Val Thr Gln Met Gly Ser Asn Met	
425 430 435 440	
aag aaa gct gtg ttt gat gag caa atg gca aaa gcg ttg aag aaa tgg	1698
Lys Lys Ala Val Phe Asp Glu Gln Met Ala Lys Ala Leu Lys Lys Trp	
445 450 455	
cac atg act gtg aag aag aag aaa ggc aaa gcg aga aag cca cca aca	1746
His Met Thr Val Lys Lys Lys Lys Gly Lys Ala Arg Lys Pro Pro Thr	
460 465 470	
gag acc ctt ggt gtt tct gac act gtc agc acc tct acc tca tcc ttt	1794
Glu Thr Leu Gly Val Ser Asp Thr Val Ser Thr Ser Thr Ser Ser Phe	
475 480 485	
cac gcc tct gga gcc act cta ctc cgc tcc aag acc act ggt cac tgc	1842
His Ala Ser Gly Ala Thr Leu Leu Arg Ser Lys Thr Thr Gly His Ser	
490 495 500	
aca gcc tct tat atg agt aat ttc gag gac caa agc atg tct gat ctt	1890
Thr Ala Ser Tyr Met Ser Asn Phe Glu Asp Gln Ser Met Ser Asp Leu	
505 510 515 520	
gaa gct gag cca tta tcc cct gaa cca ata gag ggg cac act ctc gtc	1938
Glu Ala Glu Pro Leu Ser Pro Glu Pro Ile Glu Gly His Thr Leu Val	
525 530 535	
agg gtt ggt gat cag aac aca gag ata gaa tat act gga gat att agt	1986
Arg Val Gly Asp Gln Asn Thr Glu Ile Glu Tyr Thr Gly Asp Ile Ser	
540 545 550	

cct gga aac caa ttc tcc ttt gtg aag aac gtt cct gct aat gat att 2034  
 Pro Gly Asn Gln Phe Ser Phe Val Lys Asn Val Pro Ala Asn Asp Ile  
 555 560 565

gac taatattcaa aatgaatgca gaacaaatcc atcatccggt ctttattttc 2087  
 Asp

tattacatgt atgccaacaa ttgcttcgcc aagtgttacc aactagggtt tctgtataag 2147

gctgtatttt agagctaaaa aaaaaaaaaa aaaaaaaaaa ctaaattact 2197

<210> 10

<211> 569

<212> PRT

<213> Arabidopsis thaliana

<400> 10

Met Ala Thr Arg Cys Phe Trp Cys Trp Thr Thr Leu Leu Phe Cys Ser  
 1 5 10 15

Gln Leu Leu Thr Gly Phe Ala Arg Ala Ser Ser Ala Gly Gly Ala Lys  
 20 25 30

Glu Lys Gly Leu Ser Gln Thr Pro Thr Trp Ala Val Ala Leu Val Cys  
 35 40 45

Thr Phe Phe Ile Leu Val Ser Val Leu Leu Glu Lys Ala Leu His Arg  
 50 55 60

Val Ala Thr Trp Leu Trp Glu Lys His Lys Asn Ser Leu Leu Glu Ala  
 65 70 75 80

Leu Glu Lys Ile Lys Ala Glu Leu Met Ile Leu Gly Phe Ile Ser Leu  
 85 90 95

Leu Leu Thr Phe Gly Glu Gln Tyr Ile Leu Lys Ile Cys Ile Pro Glu  
 100 105 110

Lys Ala Ala Ala Ser Met Leu Pro Cys Pro Ala Pro Ser Thr His Asp  
 115 120 125

Gln Asp Lys Thr His Arg Arg Arg Leu Ala Ala Ala Thr Thr Ser Ser  
 130 135 140

Arg Cys Asp Glu Gly His Glu Pro Leu Ile Pro Ala Thr Gly Leu His  
 145 150 155 160

Gln Leu His Ile Leu Leu Phe Phe Met Ala Ala Phe His Ile Leu Tyr  
 165 170 175

Ser Phe Ile Thr Met Met Leu Gly Arg Leu Lys Ile Arg Gly Trp Lys  
 180 185 190



Lys Trp Glu Gln Glu Thr Cys Ser His Asp Tyr Glu Phe Ser Ile Asp  
 195 200 205  
 Pro Ser Arg Phe Arg Leu Thr His Glu Thr Ser Phe Val Arg Gln His  
 210 215 220  
 Ser Ser Phe Trp Thr Lys Ile Pro Phe Phe Phe Tyr Ala Gly Cys Phe  
 225 230 235 240  
 Leu Gln Gln Phe Phe Arg Ser Val Gly Arg Thr Asp Tyr Leu Thr Leu  
 245 250 255  
 Arg His Gly Phe Ile Ala Ala His Leu Ala Pro Gly Arg Lys Phe Asp  
 260 265 270  
 Phe Gln Lys Tyr Ile Lys Arg Ser Leu Glu Asp Asp Phe Lys Val Val  
 275 280 285  
 Val Gly Ile Ser Pro Leu Leu Trp Ala Ser Phe Val Ile Phe Leu Leu  
 290 295 300  
 Leu Asn Val Asn Gly Trp Glu Ala Leu Phe Trp Ala Ser Ile Leu Pro  
 305 310 315 320  
 Val Leu Ile Ile Leu Ala Val Ser Thr Lys Leu Gln Ala Ile Leu Thr  
 325 330 335  
 Arg Met Ala Leu Gly Ile Thr Glu Arg His Ala Val Val Gln Gly Ile  
 340 345 350  
 Pro Leu Val His Gly Ser Asp Lys Tyr Phe Trp Phe Asn Arg Pro Gln  
 355 360 365  
 Leu Leu Leu His Leu Leu His Phe Ala Leu Phe Gln Asn Ala Phe Gln  
 370 375 380  
 Leu Thr Tyr Phe Phe Trp Val Trp Tyr Ser Phe Gly Leu Lys Ser Cys  
 385 390 395 400  
 Phe His Thr Asp Phe Lys Leu Val Ile Val Lys Leu Ser Leu Gly Val  
 405 410 415  
 Gly Ala Leu Ile Leu Cys Ser Tyr Ile Thr Leu Pro Leu Tyr Ala Leu  
 420 425 430  
 Val Thr Gln Met Gly Ser Asn Met Lys Lys Ala Val Phe Asp Glu Gln  
 435 440 445  
 Met Ala Lys Ala Leu Lys Lys Trp His Met Thr Val Lys Lys Lys Lys  
 450 455 460  
 Gly Lys Ala Arg Lys Pro Pro Thr Glu Thr Leu Gly Val Ser Asp Thr  
 465 470 475 480  
 Val Ser Thr Ser Thr Ser Ser Phe His Ala Ser Gly Ala Thr Leu Leu

485                      490                      495  
 Arg Ser Lys Thr Thr Gly His Ser Thr Ala Ser Tyr Met Ser Asn Phe  
                          500                      505                      510  
 Glu Asp Gln Ser Met Ser Asp Leu Glu Ala Glu Pro Leu Ser Pro Glu  
                          515                      520                      525  
 Pro Ile Glu Gly His Thr Leu Val Arg Val Gly Asp Gln Asn Thr Glu  
                          530                      535                      540  
 Ile Glu Tyr Thr Gly Asp Ile Ser Pro Gly Asn Gln Phe Ser Phe Val  
                          545                      550                      555                      560  
 Lys Asn Val Pro Ala Asn Asp Ile Asp  
    565

<210> 11  
 <211> 1935  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (89) .. (1807)

<220>  
 <221> misc\_feature  
 <222> (269) .. (307)  
 <223> location of the amino acid sequence set forth in  
                          SEQ ID No:1

<220>  
 <221> misc\_feature  
 <222> (1370) .. (1411)  
 <223> location of the amino acid sequence set forth in  
                          SEQ ID No:2

<400> 11  
 cagtgtgagt aatttagtaa aaagacaaga tctctgggtct ggaattagaa gaatcttatt 60  
 tgggtttttt tcttaggatt aagctcta atg gca gat caa gta aaa gag cgg 112  
    Met Ala Asp Gln Val Lys Glu Arg  
    1                      5  
 act tta gag gag acc tct acg tgg gca gtt gct gtt gtt tgc ttt gtc 160  
 Thr Leu Glu Glu Thr Ser Thr Trp Ala Val Ala Val Val Cys Phe Val  
                          10                      15                      20  
 tta ctc ttt att tcg att gtc ctc gaa cat tct att cac aaa att gga 208  
 Leu Leu Phe Ile Ser Ile Val Leu Glu His Ser Ile His Lys Ile Gly  
                          25                      30                      35                      40  
 acc tgg ttt aaa aag aag cac aag cag gct ctt ttt gaa gct ctt gaa 256

Thr Trp Phe Lys Lys Lys His Lys Gln Ala Leu Phe Glu Ala Leu Glu	
45 50 55	
aag gtc aaa gca gag ctt atg ctg ttg gga ttc ata tca cta cta ctc	304
Lys Val Lys Ala Glu Leu Met Leu Leu Gly Phe Ile Ser Leu Leu Leu	
60 65 70	
aca att gga caa aca cca atc tca aat atc tgc atc tcc cag aaa gtt	352
Thr Ile Gly Gln Thr Pro Ile Ser Asn Ile Cys Ile Ser Gln Lys Val	
75 80 85	
gcg tca aca atg cac cct tgc agt gct gct gaa gaa gct aaa aaa tac	400
Ala Ser Thr Met His Pro Cys Ser Ala Ala Glu Glu Ala Lys Lys Tyr	
90 95 100	
ggc aag aaa gac gcc gga aag aaa gat gat gga gat gga gat aaa ccc	448
Gly Lys Lys Asp Ala Gly Lys Lys Asp Asp Gly Asp Gly Asp Lys Pro	
105 110 115 120	
ggt cga aga ctt ctt ctt gag tta gct gaa tct tat atc cat aga aga	496
Gly Arg Arg Leu Leu Leu Glu Leu Ala Glu Ser Tyr Ile His Arg Arg	
125 130 135	
agt tta gcc acc aaa ggc tat gac aaa tgt gca gag aag ggg aaa gtg	544
Ser Leu Ala Thr Lys Gly Tyr Asp Lys Cys Ala Glu Lys Gly Lys Val	
140 145 150	
gct ttt gta tct gct tat gga atc cac cag ctg cat ata ttc atc ttc	592
Ala Phe Val Ser Ala Tyr Gly Ile His Gln Leu His Ile Phe Ile Phe	
155 160 165	
gtg ctc gcg gtt gtt cat gtt gtt tac tgc att gtt act tat gct ttc	640
Val Leu Ala Val Val His Val Val Tyr Cys Ile Val Thr Tyr Ala Phe	
170 175 180	
gga aag atc aag atg agg acg tgg aag tgg tgg gag gaa gag aca aag	688
Gly Lys Ile Lys Met Arg Thr Trp Lys Ser Trp Glu Glu Glu Thr Lys	
185 190 195 200	
aca ata gag tat cag tat tcc aac gat cct gag agg ttc agg ttt gcg	736
Thr Ile Glu Tyr Gln Tyr Ser Asn Asp Pro Glu Arg Phe Arg Phe Ala	
205 210 215	
agg gac aca tct ttt ggg aga aga cat ctc aat ttc tgg agc aag acg	784
Arg Asp Thr Ser Phe Gly Arg Arg His Leu Asn Phe Trp Ser Lys Thr	
220 225 230	
aga gtc aca cta tgg att gtt tgt ttt ttt aga cag ttc ttt gga tct	832
Arg Val Thr Leu Trp Ile Val Cys Phe Phe Arg Gln Phe Phe Gly Ser	
235 240 245	
gtc acc aaa gtt gat tac tta gca cta aga cat ggt ttc atc atg gcg	880
Val Thr Lys Val Asp Tyr Leu Ala Leu Arg His Gly Phe Ile Met Ala	
250 255 260	



ggc tca tct cca atc cat ctc ctt cac aat ttc aat aac cgg agc gtt 1600  
 Gly Ser Ser Pro Ile His Leu Leu His Asn Phe Asn Asn Arg Ser Val  
 490 495 500

gaa aat tac cca agt tct cct tct cct aga tac tct ggt cat ggt cat 1648  
 Glu Asn Tyr Pro Ser Ser Pro Ser Pro Arg Tyr Ser Gly His Gly His  
 505 510 515 520

cat gaa cac caa ttt tgg gat cct gag tct caa cac caa gaa gct gaa 1696  
 His Glu His Gln Phe Trp Asp Pro Glu Ser Gln His Gln Glu Ala Glu  
 525 530 535

act tcc aca cat cat tct ctt gcg cat gaa agc tca gaa cct gtt ctt 1744  
 Thr Ser Thr His His Ser Leu Ala His Glu Ser Ser Glu Pro Val Leu  
 540 545 550

gca tct gtg gaa ctt cct cct ata agg act agc aaa agc tta aga gat 1792  
 Ala Ser Val Glu Leu Pro Pro Ile Arg Thr Ser Lys Ser Leu Arg Asp  
 555 560 565

ttt tct ttt aag aaa tgatgattct tgtttgctat atttgatttc gtacagtggg 1847  
 Phe Ser Phe Lys Lys  
 570

aattttgtca tatgaaaata atttcttgta cattactagt ttgataagaa ataaccatat 1907  
 ctatatggat acaaaaaaaaaa aaaaaaaaaa 1935

<210> 12

<211> 573

<212> PRT

<213> Arabidopsis thaliana

<400> 12

Met Ala Asp Gln Val Lys Glu Arg Thr Leu Glu Glu Thr Ser Thr Trp  
 1 5 10 15

Ala Val Ala Val Val Cys Phe Val Leu Leu Phe Ile Ser Ile Val Leu  
 20 25 30

Glu His Ser Ile His Lys Ile Gly Thr Trp Phe Lys Lys Lys His Lys  
 35 40 45

Gln Ala Leu Phe Glu Ala Leu Glu Lys Val Lys Ala Glu Leu Met Leu  
 50 55 60

Leu Gly Phe Ile Ser Leu Leu Leu Thr Ile Gly Gln Thr Pro Ile Ser  
 65 70 75 80

Asn Ile Cys Ile Ser Gln Lys Val Ala Ser Thr Met His Pro Cys Ser  
 85 90 95

Ala Ala Glu Glu Ala Lys Lys Tyr Gly Lys Lys Asp Ala Gly Lys Lys

100	105	110
Asp Asp Gly Asp Gly Asp Lys Pro Gly Arg Arg Leu Leu Leu Glu Leu 115 120 125		
Ala Glu Ser Tyr Ile His Arg Arg Ser Leu Ala Thr Lys Gly Tyr Asp 130 135 140		
Lys Cys Ala Glu Lys Gly Lys Val Ala Phe Val Ser Ala Tyr Gly Ile 145 150 155 160		
His Gln Leu His Ile Phe Ile Phe Val Leu Ala Val Val His Val Val 165 170 175		
Tyr Cys Ile Val Thr Tyr Ala Phe Gly Lys Ile Lys Met Arg Thr Trp 180 185 190		
Lys Ser Trp Glu Glu Glu Thr Lys Thr Ile Glu Tyr Gln Tyr Ser Asn 195 200 205		
Asp Pro Glu Arg Phe Arg Phe Ala Arg Asp Thr Ser Phe Gly Arg Arg 210 215 220		
His Leu Asn Phe Trp Ser Lys Thr Arg Val Thr Leu Trp Ile Val Cys 225 230 235 240		
Phe Phe Arg Gln Phe Phe Gly Ser Val Thr Lys Val Asp Tyr Leu Ala 245 250 255		
Leu Arg His Gly Phe Ile Met Ala His Phe Ala Pro Gly Asn Glu Ser 260 265 270		
Arg Phe Asp Phe Arg Lys Tyr Ile Gln Arg Ser Leu Glu Lys Asp Phe 275 280 285		
Lys Thr Val Val Glu Ile Ser Pro Val Ile Trp Phe Val Ala Val Leu 290 295 300		
Phe Leu Leu Thr Asn Ser Tyr Gly Leu Arg Ser Tyr Leu Trp Leu Pro 305 310 315 320		
Phe Ile Pro Leu Val Val Ile Leu Ile Val Gly Thr Lys Leu Glu Val 325 330 335		
Ile Ile Thr Lys Leu Gly Leu Arg Ile Gln Glu Lys Gly Asp Val Val 340 345 350		
Arg Gly Ala Pro Val Val Gln Pro Gly Asp Asp Leu Phe Trp Phe Gly 355 360 365		
Lys Pro Arg Phe Ile Leu Phe Leu Ile His Leu Val Leu Phe Thr Asn 370 375 380		
Ala Phe Gln Leu Ala Phe Phe Ala Trp Ser Thr Tyr Glu Phe Asn Leu 385 390 395 400		

```

Asn Asn Cys Phe His Glu Ser Thr Ala Asp Val Val Ile Arg Leu Val
      405                                410                        415

Val Gly Ala Val Val Gln Ile Leu Cys Ser Tyr Val Thr Leu Pro Leu
      420                                425                        430

Tyr Ala Leu Val Thr Gln Met Gly Ser Lys Met Lys Pro Thr Val Phe
      435                                440                        445

Asn Asp Arg Val Ala Thr Ala Leu Lys Lys Trp His His Thr Ala Lys
      450                                455                        460

Asn Glu Thr Lys His Gly Arg His Ser Gly Ser Asn Thr Pro Phe Ser
465                                470                        475                        480

Ser Arg Pro Thr Thr Pro Thr His Gly Ser Ser Pro Ile His Leu Leu
      485                                490                        495

His Asn Phe Asn Asn Arg Ser Val Glu Asn Tyr Pro Ser Ser Pro Ser
      500                                505                        510

Pro Arg Tyr Ser Gly His Gly His His Glu His Gln Phe Trp Asp Pro
      515                                520                        525

Glu Ser Gln His Gln Glu Ala Glu Thr Ser Thr His His Ser Leu Ala
      530                                535                        540

His Glu Ser Ser Glu Pro Val Leu Ala Ser Val Glu Leu Pro Pro Ile
545                                550                        555                        560

Arg Thr Ser Lys Ser Leu Arg Asp Phe Ser Phe Lys Lys
      565                                570

```

&lt;210&gt; 13

&lt;211&gt; 1811

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (56)..(1633)

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (236)..(274)

<223> location of the amino acid sequence set forth in  
SEQ ID No:1

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (1328)..(1369)

<223> location of the amino acid sequence set forth in  
SEQ ID No:2

&lt;400&gt; 13

```

gttcccagat tcattctttac ttattgtcta aattctctct ggtgtgagaa gtaaa atg 58
                                     Met
                                     1

ggt cac gga gga gaa ggg atg tgg ctt gaa ttc act ccg acg tgg gtc 106
Gly His Gly Gly Glu Gly Met Ser Leu Glu Phe Thr Pro Thr Trp Val
      5                      10                      15

gtc gcc gga gtt tgt acg gtc atc gtc gcg att tca ctg gcg gtg gag 154
Val Ala Gly Val Cys Thr Val Ile Val Ala Ile Ser Leu Ala Val Glu
      20                      25                      30

cgt ttg ctt cac tat ttc ggt act gtt ctt aag aag aag aag caa aaa 202
Arg Leu Leu His Tyr Phe Gly Thr Val Leu Lys Lys Lys Lys Gln Lys
      35                      40                      45

ccc ctt tac gaa gcc ctt caa aag gtt aaa gaa gag ctg atg ttg tta 250
Pro Leu Tyr Glu Ala Leu Gln Lys Val Lys Glu Glu Leu Met Leu Leu
      50                      55                      60                      65

ggg ttt ata tgg ctg tta ctg acg gta ttc caa ggg ctc att tcc aaa 298
Gly Phe Ile Ser Leu Leu Leu Thr Val Phe Gln Gly Leu Ile Ser Lys
      70                      75                      80

ttc tgt gtg aaa gaa aat gtg ctt atg cat atg ctt cca tgt tct ctc 346
Phe Cys Val Lys Glu Asn Val Leu Met His Met Leu Pro Cys Ser Leu
      85                      90                      95

gat tca aga cga gaa gct ggg gca agt gaa cat aaa aac gtt aca gca 394
Asp Ser Arg Arg Glu Ala Gly Ala Ser Glu His Lys Asn Val Thr Ala
      100                      105                      110

aaa gaa cat ttt cag act ttt tta cct att gtt gga acc act agg cgt 442
Lys Glu His Phe Gln Thr Phe Leu Pro Ile Val Gly Thr Thr Arg Arg
      115                      120                      125

cta ctt gct gaa cat gct gct gtg caa gtt ggt tac tgt agc gaa aag 490
Leu Leu Ala Glu His Ala Ala Val Gln Val Gly Tyr Cys Ser Glu Lys
      130                      135                      140                      145

ggt aaa gta cca ttg ctt tgg ctt gag gca ttg cac cat cta cat att 538
Gly Lys Val Pro Leu Leu Ser Leu Glu Ala Leu His His Leu His Ile
      150                      155                      160

ttc atc ttc gtc ctc gcc ata tcc cat gtg aca ttc tgt gtc ctt acc 586
Phe Ile Phe Val Leu Ala Ile Ser His Val Thr Phe Cys Val Leu Thr
      165                      170                      175

gtg att ttt gga agc aca agg att cac caa tgg aag aaa tgg gag gat 634
Val Ile Phe Gly Ser Thr Arg Ile His Gln Trp Lys Lys Trp Glu Asp
      180                      185                      190

tgg atc gca gat gag aag ttt gac ccc gaa aca gct ctc agg aaa aga 682

```



Ser	Ile	Ala	Asp	Glu	Lys	Phe	Asp	Pro	Glu	Thr	Ala	Leu	Arg	Lys	Arg		
195						200					205						
agg	gtc	act	cat	gta	cac	aac	cat	gct	ttt	att	aaa	gag	cat	ttt	ctt	730	
Arg	Val	Thr	His	Val	His	Asn	His	Ala	Phe	Ile	Lys	Glu	His	Phe	Leu		
210					215				220					225			
ggg	att	ggc	aaa	gat	tca	gtc	atc	ctc	gga	tgg	acg	caa	tcc	ttt	ctc	778	
Gly	Ile	Gly	Lys	Asp	Ser	Val	Ile	Leu	Gly	Trp	Thr	Gln	Ser	Phe	Leu		
				230					235					240			
aag	caa	ttc	tat	gat	tct	gtg	acg	aaa	tca	gat	tac	gtg	act	tta	cgt	826	
Lys	Gln	Phe	Tyr	Asp	Ser	Val	Thr	Lys	Ser	Asp	Tyr	Val	Thr	Leu	Arg		
			245					250					255				
ctt	ggg	ttc	att	atg	aca	cat	tgt	aag	gga	aac	ccc	aag	ctt	aat	ttc	874	
Leu	Gly	Phe	Ile	Met	Thr	His	Cys	Lys	Gly	Asn	Pro	Lys	Leu	Asn	Phe		
			260				265				270						
cac	aag	tat	atg	atg	cgc	gct	yta	gag	gat	gat	ttc	aaa	caa	gtt	gtt	922	
His	Lys	Tyr	Met	Met	Arg	Ala	Xaa	Glu	Asp	Asp	Phe	Lys	Gln	Val	Val		
			275			280					285						
ggg	att	agt	tgg	tat	ctt	tgg	atc	ttt	gtc	gtc	atc	ttt	ttg	ctg	cta	970	
Gly	Ile	Ser	Trp	Tyr	Leu	Trp	Ile	Phe	Val	Val	Ile	Phe	Leu	Leu	Leu		
290					295				300					305			
aat	gtt	aac	gga	tgg	cac	aca	tat	ttc	tgg	ata	gca	ttt	att	ccc	ttt	1018	
Asn	Val	Asn	Gly	Trp	His	Thr	Tyr	Phe	Trp	Ile	Ala	Phe	Ile	Pro	Phe		
					310				315					320			
gct	ttg	ctt	ctt	gct	gtg	gga	aca	aag	ttg	gag	cat	gtg	att	gca	cag	1066	
Ala	Leu	Leu	Leu	Ala	Val	Gly	Thr	Lys	Leu	Glu	His	Val	Ile	Ala	Gln		
			325					330					335				
tta	gct	cat	gaa	gtt	gca	gag	aaa	cat	gta	gcc	att	gaa	gga	gac	tta	1114	
Leu	Ala	His	Glu	Val	Ala	Glu	Lys	His	Val	Ala	Ile	Glu	Gly	Asp	Leu		
			340				345					350					
gtg	gtg	aaa	ccc	tca	gat	gag	cat	ttc	tgg	ttc	agc	aaa	cct	caa	att	1162	
Val	Val	Lys	Pro	Ser	Asp	Glu	His	Phe	Trp	Phe	Ser	Lys	Pro	Gln	Ile		
			355			360					365						
gtt	ctc	tac	ttg	atc	cat	ttt	atc	ctc	ttc	cag	aat	gct	ttt	gag	att	1210	
Val	Leu	Tyr	Leu	Ile	His	Phe	Ile	Leu	Phe	Gln	Asn	Ala	Phe	Glu	Ile		
			370			375				380				385			
gcg	ttt	ttc	ttt	tgg	att	tgg	gtt	aca	tac	ggc	ttc	gac	tcg	tgc	att	1258	
Ala	Phe	Phe	Phe	Trp	Ile	Trp	Val	Thr	Tyr	Gly	Phe	Asp	Ser	Cys	Ile		
				390					395					400			
atg	gga	cag	gtg	aga	tac	att	gtt	cca	aga	ttg	gtt	atc	ggg	gtc	ttc	1306	
Met	Gly	Gln	Val	Arg	Tyr	Ile	Val	Pro	Arg	Leu	Val	Ile	Gly	Val	Phe		
			405					410					415				

att caa gtg ctt tgc agt tac agt aca ctg cct ctt tac gcc atc gtc 1354  
 Ile Gln Val Leu Cys Ser Tyr Ser Thr Leu Pro Leu Tyr Ala Ile Val  
 420 425 430

tca cag atg gga agt agc ttc aag aaa gct ata ttc gag gag aat gtg 1402  
 Ser Gln Met Gly Ser Ser Phe Lys Lys Ala Ile Phe Glu Glu Asn Val  
 435 440 445

cag gtt ggt ctt gtt ggt tgg gca cag aaa gtg aaa caa aag aga gac 1450  
 Gln Val Gly Leu Val Gly Trp Ala Gln Lys Val Lys Gln Lys Arg Asp  
 450 455 460 465

cta aaa gct gca gct agt aat gga aac gaa gga agc tct cag gct ggt 1498  
 Leu Lys Ala Ala Ala Ser Asn Gly Asn Glu Gly Ser Ser Gln Ala Gly  
 470 475 480

cct ggt cct gat tct ggt tct ggt tct gct cct gct gct ggt cct ggt 1546  
 Pro Gly Pro Asp Ser Gly Ser Gly Ser Ala Pro Ala Ala Gly Pro Gly  
 485 490 495

gca ggt ttt gca gga att cag ctc agc aga gta aca aga aac aac gca 1594  
 Ala Gly Phe Ala Gly Ile Gln Leu Ser Arg Val Thr Arg Asn Asn Ala  
 500 505 510

ggg gac aca aac aat gag att aca cct gat cat aac aac tgagcagaga 1643  
 Gly Asp Thr Asn Asn Glu Ile Thr Pro Asp His Asn Asn  
 515 520 525

tattatcttt tccatttaga ggatcatcat cagatttttag cttcaagggtc cggttttgtg 1703

gtttatacat aagttatagt gacttgattt ttttgttttg ttacaaagtt accatctttg 1763

gattagaatt gggaaattga atctgtttgt aaaaaaaaaa aaaaaaaaaa 1811

<210> 14  
 <211> 526  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 14  
 Met Gly His Gly Gly Glu Gly Met Ser Leu Glu Phe Thr Pro Thr Trp  
 1 5 10 15

Val Val Ala Gly Val Cys Thr Val Ile Val Ala Ile Ser Leu Ala Val  
 20 25 30

Glu Arg Leu Leu His Tyr Phe Gly Thr Val Leu Lys Lys Lys Lys Gln  
 35 40 45

Lys Pro Leu Tyr Glu Ala Leu Gln Lys Val Lys Glu Glu Leu Met Leu  
 50 55 60

Leu Gly Phe Ile Ser Leu Leu Leu Thr Val Phe Gln Gly Leu Ile Ser  
 65 70 75 80

Lys Phe Cys Val Lys Glu Asn Val Leu Met His Met Leu Pro Cys Ser  
                     85                    90                    95  
 Leu Asp Ser Arg Arg Glu Ala Gly Ala Ser Glu His Lys Asn Val Thr  
                     100                    105                    110  
 Ala Lys Glu His Phe Gln Thr Phe Leu Pro Ile Val Gly Thr Thr Arg  
                     115                    120                    125  
 Arg Leu Leu Ala Glu His Ala Ala Val Gln Val Gly Tyr Cys Ser Glu  
                     130                    135                    140  
 Lys Gly Lys Val Pro Leu Leu Ser Leu Glu Ala Leu His His Leu His  
                     145                    150                    155                    160  
 Ile Phe Ile Phe Val Leu Ala Ile Ser His Val Thr Phe Cys Val Leu  
                     165                    170                    175  
 Thr Val Ile Phe Gly Ser Thr Arg Ile His Gln Trp Lys Lys Trp Glu  
                     180                    185                    190  
 Asp Ser Ile Ala Asp Glu Lys Phe Asp Pro Glu Thr Ala Leu Arg Lys  
                     195                    200                    205  
 Arg Arg Val Thr His Val His Asn His Ala Phe Ile Lys Glu His Phe  
                     210                    215                    220  
 Leu Gly Ile Gly Lys Asp Ser Val Ile Leu Gly Trp Thr Gln Ser Phe  
                     225                    230                    235                    240  
 Leu Lys Gln Phe Tyr Asp Ser Val Thr Lys Ser Asp Tyr Val Thr Leu  
                     245                    250                    255  
 Arg Leu Gly Phe Ile Met Thr His Cys Lys Gly Asn Pro Lys Leu Asn  
                     260                    265                    270  
 Phe His Lys Tyr Met Met Arg Ala Xaa Glu Asp Asp Phe Lys Gln Val  
                     275                    280                    285  
 Val Gly Ile Ser Trp Tyr Leu Trp Ile Phe Val Val Ile Phe Leu Leu  
                     290                    295                    300  
 Leu Asn Val Asn Gly Trp His Thr Tyr Phe Trp Ile Ala Phe Ile Pro  
                     305                    310                    315                    320  
 Phe Ala Leu Leu Leu Ala Val Gly Thr Lys Leu Glu His Val Ile Ala  
                     325                    330                    335  
 Gln Leu Ala His Glu Val Ala Glu Lys His Val Ala Ile Glu Gly Asp  
                     340                    345                    350  
 Leu Val Val Lys Pro Ser Asp Glu His Phe Trp Phe Ser Lys Pro Gln  
                     355                    360                    365

- 32 -

Ile Val Leu Tyr Leu Ile His Phe Ile Leu Phe Gln Asn Ala Phe Glu  
 370 375 380  
 Ile Ala Phe Phe Phe Trp Ile Trp Val Thr Tyr Gly Phe Asp Ser Cys  
 385 390 395 400  
 Ile Met Gly Gln Val Arg Tyr Ile Val Pro Arg Leu Val Ile Gly Val  
 405 410 415  
 Phe Ile Gln Val Leu Cys Ser Tyr Ser Thr Leu Pro Leu Tyr Ala Ile  
 420 425 430  
 Val Ser Gln Met Gly Ser Ser Phe Lys Lys Ala Ile Phe Glu Glu Asn  
 435 440 445  
 Val Gln Val Gly Leu Val Gly Trp Ala Gln Lys Val Lys Gln Lys Arg  
 450 455 460  
 Asp Leu Lys Ala Ala Ala Ser Asn Gly Asn Glu Gly Ser Ser Gln Ala  
 465 470 475 480  
 Gly Pro Gly Pro Asp Ser Gly Ser Gly Ser Ala Pro Ala Ala Gly Pro  
 485 490 495  
 Gly Ala Gly Phe Ala Gly Ile Gln Leu Ser Arg Val Thr Arg Asn Asn  
 500 505 510  
 Ala Gly Asp Thr Asn Asn Glu Ile Thr Pro Asp His Asn Asn  
 515 520 525

<210> 15  
 <211> 1782  
 <212> DNA  
 <213> Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (1)..(1779)

<220>  
 <221> misc\_feature  
 <222> (274)..(313)  
 <223> location of the amino acid sequence set forth in  
 SEQ ID No:1

<220>  
 <221> misc\_feature  
 <222> (1327)..(1370)  
 <223> location of the amino acid sequence set forth in  
 SEQ ID No: 2

<400> 15  
 atg gga atc atc gac ggt tct ttg ctt cgg cgg ttg att tgt ctc tgt 48  
 Met Gly Ile Ile Asp Gly Ser Leu Leu Arg Arg Leu Ile Cys Leu Cys

1	5	10	15	
ctc tgg tgt ctt ctc ggt gga gga gtg acg gtg gtt acg gcg gag gat	96			
Leu Trp Cys Leu Leu Gly Gly Gly Val Thr Val Val Thr Ala Glu Asp				
20 25 30				
gag aag aaa gtg gta cat aaa cag ctt aat caa act ccg act tgg gct	144			
Glu Lys Lys Val Val His Lys Gln Leu Asn Gln Thr Pro Thr Trp Ala				
35 40 45				
gtt gct gct gtt tgt act ttc ttc atc gtt gtt tct gtt ctt ctt gaa	192			
Val Ala Ala Val Cys Thr Phe Phe Ile Val Val Ser Val Leu Leu Glu				
50 55 60				
aaa ctt ctt cac aaa gtt gga aag gtt cta tgg gat cgg cac aag aca	240			
Lys Leu Leu His Lys Val Gly Lys Val Leu Trp Asp Arg His Lys Thr				
65 70 75 80				
gct ctt ctt gac gct ttg gag aag atc aaa gca gag ctg atg gtt ctt	288			
Ala Leu Leu Asp Ala Leu Glu Lys Ile Lys Ala Glu Leu Met Val Leu				
85 90 95				
gga ttc atc tct ttg ctt ctg aca ttt gga caa acc tac att ttg gat	336			
Gly Phe Ile Ser Leu Leu Leu Thr Phe Gly Gln Thr Tyr Ile Leu Asp				
100 105 110				
att tgt atc cct tca cat gtt gct cgt acg atg ctc ccg tgt cct gct	384			
Ile Cys Ile Pro Ser His Val Ala Arg Thr Met Leu Pro Cys Pro Ala				
115 120 125				
cct aac ttg aaa aag gag gat gat gac aat ggt gaa agt cac agg aga	432			
Pro Asn Leu Lys Lys Glu Asp Asp Asp Asn Gly Glu Ser His Arg Arg				
130 135 140				
ctc ttg tgc ttt gag cac aga ttt tta tct gga ggt gaa gca tct ccc	480			
Leu Leu Ser Phe Glu His Arg Phe Leu Ser Gly Gly Glu Ala Ser Pro				
145 150 155 160				
act aaa tgc acg aag gag ggt tat gta gag ctt atc tct gcc gag gca	528			
Thr Lys Cys Thr Lys Glu Gly Tyr Val Glu Leu Ile Ser Ala Glu Ala				
165 170 175				
ctc cat cag ttg cac atc ctt ata ttc ttc tta gcc att ttc cac gtt	576			
Leu His Gln Leu His Ile Leu Ile Phe Phe Leu Ala Ile Phe His Val				
180 185 190				
ctt tac agc ttc tta act atg atg ctt gga agg ttg aag att cgc gga	624			
Leu Tyr Ser Phe Leu Thr Met Met Leu Gly Arg Leu Lys Ile Arg Gly				
195 200 205				
tgg aag cat tgg gag aat gag aca tca tcc cat aat tac gag ttt tca	672			
Trp Lys His Trp Glu Asn Glu Thr Ser Ser His Asn Tyr Glu Phe Ser				
210 215 220				
aca gac act tcc aga ttc agg cta act cat gaa aca tct ttt gtg aga	720			

Thr	Asp	Thr	Ser	Arg	Phe	Arg	Leu	Thr	His	Glu	Thr	Ser	Phe	Val	Arg		
225					230					235					240		
gcg	cac	acc	agt	ttc	tgg	acc	cgg	att	cca	ttc	ttt	ttc	tat	gtt	gga	768	
Ala	His	Thr	Ser	Phe	Trp	Thr	Arg	Ile	Pro	Phe	Phe	Phe	Tyr	Val	Gly		
				245				250						255			
tgc	ttt	ttc	aga	cag	ttt	ttc	aga	tcc	gtt	ggg	aga	act	gac	tat	ttg	816	
Cys	Phe	Phe	Arg	Gln	Phe	Phe	Arg	Ser	Val	Gly	Arg	Thr	Asp	Tyr	Leu		
			260					265					270				
aca	ttg	aga	aat	ggg	ttc	atc	gct	gtt	cat	tta	gct	cca	gga	agt	caa	864	
Thr	Leu	Arg	Asn	Gly	Phe	Ile	Ala	Val	His	Leu	Ala	Pro	Gly	Ser	Gln		
			275				280					285					
ttt	aac	ttc	caa	aaa	tac	att	aaa	aga	tcg	ttg	gag	gat	gat	ttc	aag	912	
Phe	Asn	Phe	Gln	Lys	Tyr	Ile	Lys	Arg	Ser	Leu	Glu	Asp	Asp	Phe	Lys		
	290					295					300						
gta	gtc	gtt	gga	gtc	agc	cct	gtc	ttg	tgg	gga	tct	ttt	gtg	cta	ttc	960	
Val	Val	Val	Gly	Val	Ser	Pro	Val	Leu	Trp	Gly	Ser	Phe	Val	Leu	Phe		
305					310					315				320			
ctc	ctc	cta	aat	att	gac	ggg	gag	tat	atg	atg	ttc	atc	ggc	act	gca	1008	
Leu	Leu	Leu	Asn	Ile	Asp	Gly	Glu	Tyr	Met	Met	Phe	Ile	Gly	Thr	Ala		
				325				330					335				
ata	cgg	gtt	att	atc	att	tta	gct	gta	ggg	aca	aag	ctt	caa	gcg	att	1056	
Ile	Pro	Val	Ile	Ile	Ile	Leu	Ala	Val	Gly	Thr	Lys	Leu	Gln	Ala	Ile		
			340				345						350				
atg	aca	agg	atg	gct	ctt	ggg	atc	aca	gat	aga	cat	gcg	gta	gtt	caa	1104	
Met	Thr	Arg	Met	Ala	Leu	Gly	Ile	Thr	Asp	Arg	His	Ala	Val	Val	Gln		
			355			360						365					
gga	atg	cgg	ctt	gta	caa	ggc	aac	gat	gag	tat	ttc	tgg	ttc	ggg	cgt	1152	
Gly	Met	Pro	Leu	Val	Gln	Gly	Asn	Asp	Glu	Tyr	Phe	Trp	Phe	Gly	Arg		
	370				375						380						
ccc	cat	ttg	att	ctc	cat	ctc	atg	cat	ttc	gcc	ttg	ttt	cag	aac	gca	1200	
Pro	His	Leu	Ile	Leu	His	Leu	Met	His	Phe	Ala	Leu	Phe	Gln	Asn	Ala		
385				390					395					400			
ttt	cag	atc	act	tat	ttc	ttc	tgg	ata	tgg	tat	tcc	ttt	gga	tca	gat	1248	
Phe	Gln	Ile	Thr	Tyr	Phe	Phe	Trp	Ile	Trp	Tyr	Ser	Phe	Gly	Ser	Asp		
				405				410					415				
tct	tgc	tac	cat	cct	aat	ttc	aag	att	gca	ctt	gta	aaa	gta	gcg	att	1296	
Ser	Cys	Tyr	His	Pro	Asn	Phe	Lys	Ile	Ala	Leu	Val	Lys	Val	Ala	Ile		
			420				425						430				
gct	tta	gga	gta	ttg	tgt	ctt	tgc	agc	tac	atc	aca	ctt	cct	ctt	tac	1344	
Ala	Leu	Gly	Val	Leu	Cys	Leu	Cys	Ser	Tyr	Ile	Thr	Leu	Pro	Leu	Tyr		
		435				440						445					

gca ctc gta act cag atg ggt tct cgg atg aaa aaa tcg gta ttc gat 1392  
 Ala Leu Val Thr Gln Met Gly Ser Arg Met Lys Lys Ser Val Phe Asp  
 450 455 460

gaa caa acg tca aaa gca ctc aag aaa tgg aga atg gca gtg aag aag 1440  
 Glu Gln Thr Ser Lys Ala Leu Lys Lys Trp Arg Met Ala Val Lys Lys  
 465 470 475 480

aag aaa ggt gtg aaa gcc act act aag aga cta ggt gga gat gga agt 1488  
 Lys Lys Gly Val Lys Ala Thr Thr Lys Arg Leu Gly Gly Asp Gly Ser  
 485 490 495

gcg agc cct acg gca tcg aca gtt agg tct act tcg tct gta cgt tca 1536  
 Ala Ser Pro Thr Ala Ser Thr Val Arg Ser Thr Ser Ser Val Arg Ser  
 500 505 510

ttg cag cgt tac aaa acc aca cca cat tcg atg aga tac gaa gga ctt 1584  
 Leu Gln Arg Tyr Lys Thr Thr Pro His Ser Met Arg Tyr Glu Gly Leu  
 515 520 525

gac cct gaa aca tcg gat ctc gac aca gat aat gaa gct ttg act cct 1632  
 Asp Pro Glu Thr Ser Asp Leu Asp Thr Asp Asn Glu Ala Leu Thr Pro  
 530 535 540

ccc aaa tct cct cca agc ttc gag ctt gtt gtg aaa gtt gaa cca aat 1680  
 Pro Lys Ser Pro Pro Ser Phe Glu Leu Val Val Lys Val Glu Pro Asn  
 545 550 555 560

aag acc aat acc ggt gag act agc cgt gac act gaa act gat tct aaa 1728  
 Lys Thr Asn Thr Gly Glu Thr Ser Arg Asp Thr Glu Thr Asp Ser Lys  
 565 570 575

gag ttc tct ttc gtc aag cct gct ccg agt aat gaa tca tct caa gac 1776  
 Glu Phe Ser Phe Val Lys Pro Ala Pro Ser Asn Glu Ser Ser Gln Asp  
 580 585 590

cgg tga 1782  
 Arg

<210> 16  
 <211> 593  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 16  
 Met Gly Ile Ile Asp Gly Ser Leu Leu Arg Arg Leu Ile Cys Leu Cys  
 1 5 10 15  
 Leu Trp Cys Leu Leu Gly Gly Gly Val Thr Val Val Thr Ala Glu Asp  
 20 25 30  
 Glu Lys Lys Val Val His Lys Gln Leu Asn Gln Thr Pro Thr Trp Ala  
 35 40 45

Val Ala Ala Val Cys Thr Phe Phe Ile Val Val Ser Val Leu Leu Glu  
 50 55 60  
 Lys Leu Leu His Lys Val Gly Lys Val Leu Trp Asp Arg His Lys Thr  
 65 70 75 80  
 Ala Leu Leu Asp Ala Leu Glu Lys Ile Lys Ala Glu Leu Met Val Leu  
 85 90 95  
 Gly Phe Ile Ser Leu Leu Leu Thr Phe Gly Gln Thr Tyr Ile Leu Asp  
 100 105 110  
 Ile Cys Ile Pro Ser His Val Ala Arg Thr Met Leu Pro Cys Pro Ala  
 115 120 125  
 Pro Asn Leu Lys Lys Glu Asp Asp Asp Asn Gly Glu Ser His Arg Arg  
 130 135 140  
 Leu Leu Ser Phe Glu His Arg Phe Leu Ser Gly Gly Glu Ala Ser Pro  
 145 150 155 160  
 Thr Lys Cys Thr Lys Glu Gly Tyr Val Glu Leu Ile Ser Ala Glu Ala  
 165 170 175  
 Leu His Gln Leu His Ile Leu Ile Phe Phe Leu Ala Ile Phe His Val  
 180 185 190  
 Leu Tyr Ser Phe Leu Thr Met Met Leu Gly Arg Leu Lys Ile Arg Gly  
 195 200 205  
 Trp Lys His Trp Glu Asn Glu Thr Ser Ser His Asn Tyr Glu Phe Ser  
 210 215 220  
 Thr Asp Thr Ser Arg Phe Arg Leu Thr His Glu Thr Ser Phe Val Arg  
 225 230 235 240  
 Ala His Thr Ser Phe Trp Thr Arg Ile Pro Phe Phe Phe Tyr Val Gly  
 245 250 255  
 Cys Phe Phe Arg Gln Phe Phe Arg Ser Val Gly Arg Thr Asp Tyr Leu  
 260 265 270  
 Thr Leu Arg Asn Gly Phe Ile Ala Val His Leu Ala Pro Gly Ser Gln  
 275 280 285  
 Phe Asn Phe Gln Lys Tyr Ile Lys Arg Ser Leu Glu Asp Asp Phe Lys  
 290 295 300  
 Val Val Val Gly Val Ser Pro Val Leu Trp Gly Ser Phe Val Leu Phe  
 305 310 315 320  
 Leu Leu Leu Asn Ile Asp Gly Glu Tyr Met Met Phe Ile Gly Thr Ala  
 325 330 335  
 Ile Pro Val Ile Ile Ile Leu Ala Val Gly Thr Lys Leu Gln Ala Ile



340	345	350
Met Thr Arg Met Ala Leu Gly Ile Thr Asp Arg His Ala Val Val Gln		
355	360	365
Gly Met Pro Leu Val Gln Gly Asn Asp Glu Tyr Phe Trp Phe Gly Arg		
370	375	380
Pro His Leu Ile Leu His Leu Met His Phe Ala Leu Phe Gln Asn Ala		
385	390	395
Phe Gln Ile Thr Tyr Phe Phe Trp Ile Trp Tyr Ser Phe Gly Ser Asp		
405	410	415
Ser Cys Tyr His Pro Asn Phe Lys Ile Ala Leu Val Lys Val Ala Ile		
420	425	430
Ala Leu Gly Val Leu Cys Leu Cys Ser Tyr Ile Thr Leu Pro Leu Tyr		
435	440	445
Ala Leu Val Thr Gln Met Gly Ser Arg Met Lys Lys Ser Val Phe Asp		
450	455	460
Glu Gln Thr Ser Lys Ala Leu Lys Lys Trp Arg Met Ala Val Lys Lys		
465	470	475
Lys Lys Gly Val Lys Ala Thr Thr Lys Arg Leu Gly Gly Asp Gly Ser		
485	490	495
Ala Ser Pro Thr Ala Ser Thr Val Arg Ser Thr Ser Ser Val Arg Ser		
500	505	510
Leu Gln Arg Tyr Lys Thr Thr Pro His Ser Met Arg Tyr Glu Gly Leu		
515	520	525
Asp Pro Glu Thr Ser Asp Leu Asp Thr Asp Asn Glu Ala Leu Thr Pro		
530	535	540
Pro Lys Ser Pro Pro Ser Phe Glu Leu Val Val Lys Val Glu Pro Asn		
545	550	555
Lys Thr Asn Thr Gly Glu Thr Ser Arg Asp Thr Glu Thr Asp Ser Lys		
565	570	575
Glu Phe Ser Phe Val Lys Pro Ala Pro Ser Asn Glu Ser Ser Gln Asp		
580	585	590

Arg

&lt;210&gt; 17

&lt;211&gt; 1629

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

<220>  
 <221> CDS  
 <222> (1)..(1626)

<220>  
 <221> misc\_feature  
 <222> (184)..(223)  
 <223> location of the amino acid sequence set forth in  
 SEQ ID No:1

<220>  
 <221> misc\_feature  
 <222> (1141)..(1183)  
 <223> location of the amino acid sequence set forth in  
 SEQ ID No:2

<400> 17  
 atg gag cat atg atg aaa gaa gga agg tct ctt gca gag acg ccg act 48  
 Met Glu His Met Met Lys Glu Gly Arg Ser Leu Ala Glu Thr Pro Thr  
 1 5 10 15

tac tct gtt gct tcg gtt gtt act gtt ttg gtc ttt gtt tgc ttt ctc 96  
 Tyr Ser Val Ala Ser Val Val Thr Val Leu Val Phe Val Cys Phe Leu  
 20 25 30

gtt gaa cgc gcc att tac aga ttt gga aag tgg tta aag aag act aga 144  
 Val Glu Arg Ala Ile Tyr Arg Phe Gly Lys Trp Leu Lys Lys Thr Arg  
 35 40 45

aga aag gca ctt ttt act tca ctt gag aaa atg aaa gag gag ttg atg 192  
 Arg Lys Ala Leu Phe Thr Ser Leu Glu Lys Met Lys Glu Glu Leu Met  
 50 55 60

ttg ctg gga ctt ata tca ctt ctg ttg tca caa agc gcg aga tgg att 240  
 Leu Leu Gly Leu Ile Ser Leu Leu Leu Ser Gln Ser Ala Arg Trp Ile  
 65 70 75 80

tca gaa atc tgt gtt aac tct tcc ctt ttc aat agt aaa ttc tac att 288  
 Ser Glu Ile Cys Val Asn Ser Ser Leu Phe Asn Ser Lys Phe Tyr Ile  
 85 90 95

tgc tct gaa gag gac tat gga atc cat aag aaa gtt ctt ctg gaa cac 336  
 Cys Ser Glu Glu Asp Tyr Gly Ile His Lys Lys Val Leu Leu Glu His  
 100 105 110

acc tct tct aca aac cag agc tcc tta cct cat cat gga ata cat gaa 384  
 Thr Ser Ser Thr Asn Gln Ser Ser Leu Pro His His Gly Ile His Glu  
 115 120 125

gcc tct cat caa tgt ggt cat ggc cgt gaa cca ttt gtg tcg tat gag 432  
 Ala Ser His Gln Cys Gly His Gly Arg Glu Pro Phe Val Ser Tyr Glu  
 130 135 140

gga ctc gag caa ctc cta aga ttc tta ttc gtc ctg ggt atc act cat 480

Gly Leu Glu Gln Leu Leu Arg Phe Leu Phe Val Leu Gly Ile Thr His	
145	150 155 160
ggt cta tac agt ggc att gcc att ggt tta gcc atg agc aag att tac	528
Val Leu Tyr Ser Gly Ile Ala Ile Gly Leu Ala Met Ser Lys Ile Tyr	
	165 170 175
agt tgg aga aaa tgg gaa gcc caa gcg atc ata atg gct gaa tca gat	576
Ser Trp Arg Lys Trp Glu Ala Gln Ala Ile Ile Met Ala Glu Ser Asp	
	180 185 190
atc cac ctt tgt ttc ctg cgg caa ttt aga ggc tcc ata cga aag tct	624
Ile His Leu Cys Phe Leu Arg Gln Phe Arg Gly Ser Ile Arg Lys Ser	
	195 200 205
gac tac ttc gca ctt cgg tta ggt ttc ctc act aaa cat aat ttg cca	672
Asp Tyr Phe Ala Leu Arg Leu Gly Phe Leu Thr Lys His Asn Leu Pro	
	210 215 220
ttt aca tac aac ttc cat atg tat atg gta cgg acg atg gaa gat gag	720
Phe Thr Tyr Asn Phe His Met Tyr Met Val Arg Thr Met Glu Asp Glu	
	225 230 235 240
ttt cat ggc att gtt gga att agc tgg cca ctt tgg gtt tac gct ata	768
Phe His Gly Ile Val Gly Ile Ser Trp Pro Leu Trp Val Tyr Ala Ile	
	245 250 255
gta tgc atc tgc ata aat gtt cat ggc ctg aat atg tac ttt tgg ata	816
Val Cys Ile Cys Ile Asn Val His Gly Leu Asn Met Tyr Phe Trp Ile	
	260 265 270
tca ttc gtt cct gcc att ctt gtc atg ttg gtt gga acc aaa ctt gag	864
Ser Phe Val Pro Ala Ile Leu Val Met Leu Val Gly Thr Lys Leu Glu	
	275 280 285
cat gtt gtc tcc aag ctt gct ctc gag gtt aag gag cag cag aca ggc	912
His Val Val Ser Lys Leu Ala Leu Glu Val Lys Glu Gln Gln Thr Gly	
	290 295 300
aca tct aat ggg gct caa gtc aaa cca cgt gat ggg ctc ttc tgg ttt	960
Thr Ser Asn Gly Ala Gln Val Lys Pro Arg Asp Gly Leu Phe Trp Phe	
	305 310 315 320
ggg aaa cca gaa att ctg cta cgg ttg ata caa ttt atc att ttt cag	1008
Gly Lys Pro Glu Ile Leu Leu Arg Leu Ile Gln Phe Ile Ile Phe Gln	
	325 330 335
aat gca ttt gaa atg gca aca ttc atc tgg ttc ttg tgg gga atc aag	1056
Asn Ala Phe Glu Met Ala Thr Phe Ile Trp Phe Leu Trp Gly Ile Lys	
	340 345 350
gaa aga tct tgc ttc atg aag aac cat gtg atg ata tca agc cgg cta	1104
Glu Arg Ser Cys Phe Met Lys Asn His Val Met Ile Ser Ser Arg Leu	
	355 360 365

att tct ggg gtt ctc gtt cag ttc tgg tgt agt tat ggc act gtg cct 1152  
 Ile Ser Gly Val Leu Val Gln Phe Trp Cys Ser Tyr Gly Thr Val Pro  
 370 375 380

ctc aat gta atc gtt act cag atg gga tct cgg cat aag aaa gct gtg 1200  
 Leu Asn Val Ile Val Thr Gln Met Gly Ser Arg His Lys Lys Ala Val  
 385 390 395 400

ata gca gag agc gta aga gac tca ctt cac agt tgg tgc aag aga gtg 1248  
 Ile Ala Glu Ser Val Arg Asp Ser Leu His Ser Trp Cys Lys Arg Val  
 405 410 415

aaa gag agg tct aag cac acg aga tca gtg tgt tcc ctt gac aca gca 1296  
 Lys Glu Arg Ser Lys His Thr Arg Ser Val Cys Ser Leu Asp Thr Ala  
 420 425 430

aca ata gac gag aga gac gag atg aca gtg ggg aca ttg tct agg agc 1344  
 Thr Ile Asp Glu Arg Asp Glu Met Thr Val Gly Thr Leu Ser Arg Ser  
 435 440 445

tca tcg atg act tca ctg aat cag att acc ata aac tcc ata gac caa 1392  
 Ser Ser Met Thr Ser Leu Asn Gln Ile Thr Ile Asn Ser Ile Asp Gln  
 450 455 460

gca gag tct ata ttc gga gca gca gct tca tcc agc agt cct caa gat 1440  
 Ala Glu Ser Ile Phe Gly Ala Ala Ala Ser Ser Ser Pro Gln Asp  
 465 470 475 480

gga tac acg tcg agg gtg gaa gaa tat ctg tct gaa aca tac aat aac 1488  
 Gly Tyr Thr Ser Arg Val Glu Glu Tyr Leu Ser Glu Thr Tyr Asn Asn  
 485 490 495

atc ggt tcg ata ccg cct tta aac gat gag att gag att gag att gaa 1536  
 Ile Gly Ser Ile Pro Pro Leu Asn Asp Glu Ile Glu Ile Glu Ile Glu  
 500 505 510

ggt gaa gaa gat aat gga ggg aga gga agt ggg agt gat gag aat aac 1584  
 Gly Glu Glu Asp Asn Gly Gly Arg Gly Ser Gly Ser Asp Glu Asn Asn  
 515 520 525

ggt gat gct gga gaa aca ctt ctt gag ttg ttt agg agg act tga 1629  
 Gly Asp Ala Gly Glu Thr Leu Leu Glu Leu Phe Arg Arg Thr  
 530 535 540

&lt;210&gt; 18

&lt;211&gt; 542

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 18

Met Glu His Met Met Lys Glu Gly Arg Ser Leu Ala Glu Thr Pro Thr  
 1 5 10 15

Tyr Ser Val Ala Ser Val Val Thr Val Leu Val Phe Val Cys Phe Leu

20					25					30					
Val	Glu	Arg	Ala	Ile	Tyr	Arg	Phe	Gly	Lys	Trp	Leu	Lys	Lys	Thr	Arg
	35						40					45			
Arg	Lys	Ala	Leu	Phe	Thr	Ser	Leu	Glu	Lys	Met	Lys	Glu	Glu	Leu	Met
	50					55					60				
Leu	Leu	Gly	Leu	Ile	Ser	Leu	Leu	Leu	Ser	Gln	Ser	Ala	Arg	Trp	Ile
	65					70					75				80
Ser	Glu	Ile	Cys	Val	Asn	Ser	Ser	Leu	Phe	Asn	Ser	Lys	Phe	Tyr	Ile
				85					90					95	
Cys	Ser	Glu	Glu	Asp	Tyr	Gly	Ile	His	Lys	Lys	Val	Leu	Leu	Glu	His
			100					105					110		
Thr	Ser	Ser	Thr	Asn	Gln	Ser	Ser	Leu	Pro	His	His	Gly	Ile	His	Glu
			115				120					125			
Ala	Ser	His	Gln	Cys	Gly	His	Gly	Arg	Glu	Pro	Phe	Val	Ser	Tyr	Glu
	130					135					140				
Gly	Leu	Glu	Gln	Leu	Leu	Arg	Phe	Leu	Phe	Val	Leu	Gly	Ile	Thr	His
	145					150					155				160
Val	Leu	Tyr	Ser	Gly	Ile	Ala	Ile	Gly	Leu	Ala	Met	Ser	Lys	Ile	Tyr
				165					170					175	
Ser	Trp	Arg	Lys	Trp	Glu	Ala	Gln	Ala	Ile	Ile	Met	Ala	Glu	Ser	Asp
			180					185					190		
Ile	His	Leu	Cys	Phe	Leu	Arg	Gln	Phe	Arg	Gly	Ser	Ile	Arg	Lys	Ser
		195					200					205			
Asp	Tyr	Phe	Ala	Leu	Arg	Leu	Gly	Phe	Leu	Thr	Lys	His	Asn	Leu	Pro
	210					215					220				
Phe	Thr	Tyr	Asn	Phe	His	Met	Tyr	Met	Val	Arg	Thr	Met	Glu	Asp	Glu
	225					230					235				240
Phe	His	Gly	Ile	Val	Gly	Ile	Ser	Trp	Pro	Leu	Trp	Val	Tyr	Ala	Ile
			245						250					255	
Val	Cys	Ile	Cys	Ile	Asn	Val	His	Gly	Leu	Asn	Met	Tyr	Phe	Trp	Ile
			260					265					270		
Ser	Phe	Val	Pro	Ala	Ile	Leu	Val	Met	Leu	Val	Gly	Thr	Lys	Leu	Glu
		275					280					285			
His	Val	Val	Ser	Lys	Leu	Ala	Leu	Glu	Val	Lys	Glu	Gln	Gln	Thr	Gly
	290					295					300				
Thr	Ser	Asn	Gly	Ala	Gln	Val	Lys	Pro	Arg	Asp	Gly	Leu	Phe	Trp	Phe
	305					310					315				320

<400> 19  
atgtcggaca aaaaaggggt

<210> 20  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 20  
atgctaccac acgcagatcg

20

<210> 21  
<211> 21  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 21  
acagagacca cctccttgga a

21

<210> 22  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 22  
cagaaacttg tctcatccct gg

22

<210> 23  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 23  
aagaactgcc tgaagaaggc

20

<210> 24

<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 24  
caccaccttc atgatgctca

20

<210> 25  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 25  
ttccagcacc ggcacaagaa

20

<210> 26  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 26  
tggacctctt catgttcgat cccatctg

28

<210> 27  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 27  
cctgacgctg ttccagaatg cgtttca

27

<210> 28  
<211> 20  
<212> DNA  
<213> Artificial Sequence



&lt;220&gt;

<223> Description of Artificial Sequence:  
Oligonucleotide

&lt;400&gt; 28

acttctgcag gtcgactcta

20

&lt;210&gt; 29

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:  
Oligonucleotide

&lt;400&gt; 29

aagatcaaga tgaggacgtg gaagtcgtgg

30

&lt;210&gt; 30

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:  
Oligonucleotide

&lt;400&gt; 30

aggctgaacc actggggcgc ctctcaccac

30

&lt;210&gt; 31

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:  
Oligonucleotide

&lt;400&gt; 31

caagtatatg atgcgcgctc tagaggatga

30

&lt;210&gt; 32

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence:

## Oligonucleotide

<400> 32  
aggtttcacc actaagtctc cttcaatggc 30

<210> 33  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 33  
gatcattcaa gacttaggct cactcatgag 30

<210> 34  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 34  
aacagcaagg aagattacaa atgatgcccc 30

<210> 35  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 35  
ggattaagat ctaatggc 18

<210> 36  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 36

caaagatctt catttcttaa aag

23

<210> 37

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 37

gcggatccat gtcggacaaa aaagg

25

<210> 38

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 38

gcggatcctc atccctggct gaagg

25

<210> 39

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 39

ggatccacca tggccacaag atg

23

<210> 40

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 40

ggatccttag tcaatatcat tagc

24

<210> 41  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 41  
gcggatccat gggtcacgga ggagaag

27

<210> 42  
<211> 27  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 42  
gcggatcctc agttgttatg atcagga

27